

**KLASIFIKASI PENDERITA *STUNTING*
DENGAN METODE *SUPPORT VECTOR MACHINE*
(STUDI KASUS: LIMA PUSKESMAS DI KOTA BANDAR LAMPUNG)**

(Skripsi)

Oleh

**JOSUA FERNANDES NABABAN
NPM 1657051019**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

**KLASIFIKASI PENDERITA *STUNTING*
DENGAN METODE *SUPPORT VECTOR MACHINE*
(STUDI KASUS: LIMA PUSKESMAS DI KOTA BANDAR LAMPUNG)**

Oleh

JOSUA FERNANDES NABABAN

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA ILMU KOMPUTER**

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

ABSTRAK

KLASIFIKASI PENDERITA STUNTING DENGAN METODE SUPPORT VECTOR MACHINE (STUDI KASUS: LIMA PUSKESMAS DI KOTA BANDAR LAMPUNG)

Oleh

JOSUA FERNANDES NABABAN

Stunting adalah kondisi apabila tinggi badan anak tidak sesuai dengan umur yang dimiliki. Kondisi ini terjadi sebagai tanda adanya masalah gizi yang kronis pada sistem pertumbuhan anak. Stunting merupakan masalah yang cukup serius terhadap kelangsungan sumber daya manusia pada suatu negara. Oleh karena itu pemerintah berfokus untuk menurunkan angka penderita stunting di Indonesia. Diagnosis pemantuan terkait stunting masih dilakukan secara langsung dengan mengukur indikator nutrisi, tetapi proses tersebut memerlukan biaya yang tidak sedikit. Berdasarkan permasalahan tersebut penelitian ini menggunakan data non-nutritional sebagai bahan pendukung penelitian untuk memudahkan proses klasifikasi terhadap penderita stunting.

Penelitian ini bertujuan memberikan model prediksi untuk klasifikasi penderita stunting, mengukur perbandingan dari hasil analisis korelasi fitur terhadap variabel yang digunakan, dan mengukur serta mengevaluasi hasil kinerja dari metode support vector machine dalam melakukan proses klasifikasi. Penelitian ini juga menggunakan data lima puskesmas di Kota Bandar Lampung dengan pembagian data 10-fold cross validation dan menggunakan tiga kernel, yaitu linear, gaussian, dan polynomial.

Hasil dari penelitian ini adalah berhasil dibuatnya model prediksi untuk klasifikasi penderita stunting yang bertujuan untuk memudahkan dalam mengetahui hasil dari prediksi. Didapatkan juga, hasil analisis korelasi fitur seluruhnya bernilai positif, dengan korelasi tertinggi sebesar 0,92. Kemudian hasil klasifikasi menggunakan metode support vector machine didapatkan hasil akurasi tertinggi, yaitu sebesar 80,8% pada kernel linear.

Kata kunci: K-Fold Cross Validation, Kernel, Klasifikasi, Stunting, Support Vector Machine

ABSTRACT

CLASSIFICATION OF STUNTING SUFFERER USING THE SUPPORT VECTOR MACHINE METHOD (CASE STUDY: FIVE HEALTH CENTERS IN BANDAR LAMPUNG)

By

JOSUA FERNANDES NABABAN

Stunting is a condition when the child's height does not fit with his/her age. This condition occurs as a sign of chronic nutritional problems in the child's growth system. Stunting is a serious problem for the sustainability of human resources in a country. Therefore the government focuses on decreasing the number of stunting sufferers in Indonesia. The monitoring of stunting diagnosis is still carried out directly by measuring nutritional indicators, but the process costs are expensive. Based on these problems, this research uses non-nutritional data as supporting material for research to facilitate the classification process for stunting sufferers. This study aims to provide a prediction model for the classification of stunting sufferers, measuring the results of feature correlation analysis against the variables used, and measuring as well as evaluating the performance of support vector machine methods in conducting the classification process. This study used some data from five health centers in Bandar Lampung city with 10-fold cross validation and using three kernels, namely linear, gaussian, and a polynomial. The result of this study is the successful creation of a prediction model for the classification of stunting sufferer that aims to ease the result of predictions. Also obtained, the result of feature correlation analysis is entirely positive, with the highest correlation is 0,92. Then the classification result using the support vector machine method obtained the highest accuracy result, 80,8% in the linear kernel.

Keywords: K-Fold Cross Validation, Kernel, Classification, Stunting, Support Vector Machine

Judul Skripsi : **KLASIFIKASI PENDERITA *STUNTING* DENGAN
METODE *SUPPORT VECTOR MACHINE*
(STUDI KASUS: LIMA PUSKESMAS
DI KOTA BANDAR LAMPUNG)**

Nama Mahasiswa : **Josua Fernandes Nababan**

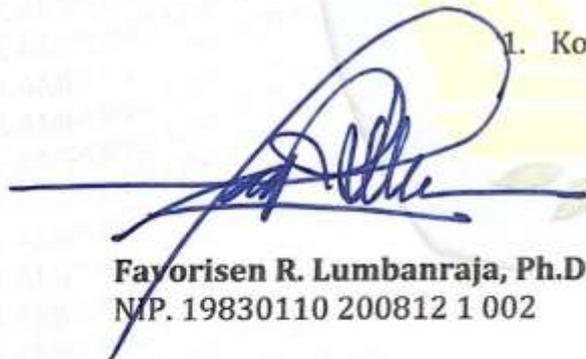
Nomor Pokok Mahasiswa : **1657051019**

Jurusan : **Ilmu Komputer**

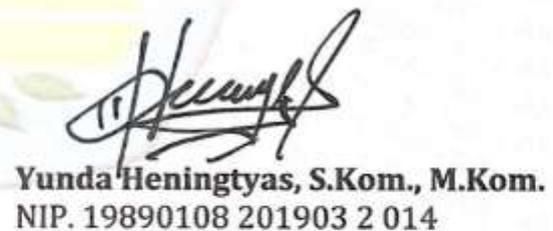
Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing



Favorisen R. Lumbanraja, Ph.D.
NIP. 19830110 200812 1 002



Yunda Heningtyas, S.Kom., M.Kom.
NIP. 19890108 201903 2 014

2. Ketua Jurusan Ilmu Komputer



Didik Kurniawan, S.Si., M.T.
NIP 19800419 200501 1 004

MENGESAHKAN

1. Tim Penguji

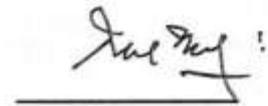
Ketua : Favorisen R. Lumbanraja, Ph.D.



Sekretaris : Yunda Heningtyas, S.Kom., M.Kom.



**Penguji
Bukan Pembimbing : Dr. Ir. Kurnia Muludi, M.S.Sc.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Supto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : 24 Juni 2021

PERNYATAAN

Saya yang bertanda tangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul “Klasifikasi Penderita *Stunting* Dengan Metode *Support Vector Machine* (Studi Kasus: Lima Puskesmas Di Kota Bandar Lampung)” merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila dikemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang saya terima.

Bandar Lampung, 24 Juni 2021



Josua Fernandes Nababan
NPM. 1657051019

RIWAYAT HIDUP



Penulis dilahirkan di Bandar Lampung pada 7 Mei 1998, sebagai anak pertama dari dua bersaudara dari ayah yang bernama Berfin Gamal Nababan dan ibu bernama Delvi Ellys Simorangkir. Penulis menyelesaikan pendidikan formal pertama kali di TK Xaverius Kalianda pada tahun 2004.

Pendidikan dasar di SD Negeri 1 Way Urang yang diselesaikan pada tahun 2010. Pendidikan menengah pertama di SMP Negeri 1 Kalianda yang diselesaikan pada tahun 2013, kemudian melanjutkan pendidikan di SMA Negeri 1 Kalianda yang diselesaikan pada tahun 2016.

Pada tahun 2016 penulis melanjutkan pendidikan dan terdaftar sebagai mahasiswa jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung. Selama menjadi mahasiswa penulis melakukan beberapa kegiatan antara lain.

1. Menjadi anggota ABACUS (Anggota Baru Computer Science) Jurusan Ilmu Komputer Universitas Lampung pada periode 2016.
2. Menjadi anggota Hubungan Masyarakat Unit Kegiatan Mahasiswa Kristen Universitas Lampung pada periode kepengurusan tahun 2018.

3. Menjadi ketua divisi Informasi dan Dasar Kepemimpinan Kristen Unit Kegiatan Mahasiswa Kristen Universitas Lampung pada periode kepengurusan tahun 2019.
4. Menjadi asisten dosen untuk beberapa mata kuliah di Jurusan Ilmu Komputer.
5. Melaksanakan kerja praktik di Otoritas Jasa Keuangan Provinsi Lampung pada tahun 2019.
6. Melaksanakan kegiatan Kuliah Kerja Nyata (KKN) di Desa Lombok, Kecamatan Lombok Seminung, Lampung Barat pada bulan Juni tahun 2019.

PERSEMBAHAN

Puji syukur kepada Tuhan Yesus Kristus atas berkat dan penyertaan-Nya telah memberikan kesehatan dan kekuatan untuk dapat menyelesaikan skripsi ini.

Skripsi ini saya persembahkan kepada:

Kedua orang tua, Bapak dan Mama yang telah memberikan kasih sayang, doa dan dukungan moral maupun material, dan yang selalu mendidik anak-anaknya menjadi pribadi yang lebih baik. Terima kasih juga kepada Mami yang telah berbahagia di surga, terima kasih telah menjadi alasan untuk selalu tetap kuat.

Keluarga Ilmu Komputer.

Fakultas Matematika dan Ilmu Pengetahuan Alam.

Almamater tercinta, Universitas Lampung.

MOTTO

“Orang-orang yang menabur dengan mencururkan air mata, akan menuai dengan bersorak-sorai. Orang yang berjalan maju dengan menangis sambil menabur benih, pasti pulang dengan sorak-sorai sambil membawa berkas-berkasnya.”

(Mazmur 126:5-6)

“Sebab Aku ini mengetahui rancangan-rancangan apa yang ada pada-Ku mengenai kamu, demikianlah firman Tuhan, yaitu rancangan damai sejahtera dan bukan rancangan kecelakaan, untuk memberikan kepadamu hari depan yang penuh harapan.” (Yeremia 29:11)

“Kamu adalah karya Tuhan yang sangat luar biasa. Jadilah dirimu sendiri dan bermanfaatlah untuk orang disekitarmu.” (Anonim)

UCAPAN TERIMA KASIH

Puji syukur kepada Tuhan Yesus Kristus, karena atas berkat dan penyertaan-Nya penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Penderita *Stunting* Dengan Metode *Support Vector Machine* (Studi Kasus: Lima Puskesmas Di Kota Bandar Lampung)”. Skripsi ini merupakan salah satu syarat untuk memperoleh gelar sarjana Ilmu Komputer di Universitas Lampung.

Dalam penyusunan skripsi ini penulis menyadari banyak mendapat bimbingan, dukungan, dan motivasi dari berbagai pihak yang dengan tulus telah membantu penulis untuk menyelesaikan skripsi ini. Penulis sangat bersyukur dan mengucapkan terima kasih sedalam-dalamnya kepada:

1. Kedua orang tua, Bapak dan Mama, serta adikku Laura Sicilia Nababan yang telah senantiasa memberikan doa, dukungan, kasih sayang yang sangat luar biasa, dan menjadi alasan untuk tetap bertahan dan menjalani hidup sampai saat ini.
2. Dedi, Mami, Kakak, Abang, dan saudara terdekat lainnya yang selalu memberikan doa, dukungan, dan kasih sayang yang sangat luar biasa, sehingga penulisan skripsi ini dapat terselesaikan dengan baik dan lancar.
3. Bapak Favorisen R. Lumbanraja, Ph.D., selaku dosen pembimbing utama atas kesediaannya dan kesabarannya untuk memberikan dukungan, bimbingan, kritik, dan saran dalam proses penyelesaian skripsi.
4. Ibu Yunda Heningtyas, S.Kom., M.Kom., selaku dosen pembimbing kedua atas kesediaannya dalam memberikan bimbingan, nasihat, juga kritik, dan saran selama proses pengerjaan skripsi.
5. Bapak Dr. Ir. Kurnia Muludi, M.S.Sc., selaku dosen pembahas skripsi yang telah memberikan saran dan masukan guna penyempurnaan penulisan skripsi.

6. Bapak Dr. Eng. Suropto Dwi Yuwono, M.T. selaku Dekan FMIPA Universitas Lampung.
7. Bapak Didik Kurniawan, S.Si., MT., selaku Ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung.
8. Ibu Astria Hijriani, S.Kom., M.Kom., selaku Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung.
9. Bapak dan Ibu Dosen Jurusan Ilmu Komputer Universitas Lampung yang telah memberikan ilmu dan pengetahuan hidup selama penulis menjadi mahasiswa.
10. Ibu Ade Nora Maela, Pak Zainudin, dan Mas Nofal yang telah memudahkan segala urusan administrasi penulis di Jurusan Ilmu Komputer.
11. Teman-teman di Jurusan Ilmu Komputer Angkatan 2016 yang menjadi teman satu angkatan selama menjalankan masa studi di Universitas Lampung.
12. Teman-teman di Unit Kegiatan Mahasiswa Kristen Universitas Lampung yang telah mengajarkan banyak hal dalam berpelayanan dan berorganisasi selama kepengurusan dan kepanitiaan.
13. Teman-teman seperbimbingan yang telah menemani dan membantu untuk sama-sama belajar dalam menyusun sampai menyelesaikan skripsi ini.
14. Teman-teman Kriskat yang telah senantiasa menjadi teman yang menemani keseharian penulis selama masa perkuliahan dengan canda dan tawa.
15. Teman-teman luar biasa Adelia Nada Efendi, Ayi Nugraha, Dwiki Prakasa, Hana Salsa Billa, Muhammad Ibrahim, Riko Rahardiansyah, dan Rehta Yunani yang selalu menjadi tempat penulis mencurahkan seluruh keluh kesah dalam segala hal, dan selalu menyemangati penulis dalam menyelesaikan skripsi ini.
16. Semua pihak yang secara langsung maupun tidak langsung yang telah membantu menyelesaikan skripsi ini.
17. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, for just being me at all times.*

Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna karena masih terbatasnya pengetahuan, pengalaman, dan kemampuan penulis. Oleh karena itu penulis sangat menghargai dan mengharapkan saran dan kritik untuk penelitian ini sebagai bahan pertimbangan untuk karya tulis yang akan datang. Semoga skripsi ini dapat bermanfaat bagi semua pihak.

Bandar Lampung, 24 Juni 2021

Josua Fernandes Nababan
NPM. 1657051019

DAFTAR ISI

Halaman

DAFTAR GAMBAR.....	xvii
DAFTAR TABEL	xviii
DAFTAR KODE PROGRAM.....	xix
I. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
II. TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu	6
2.2 Uraian Landasan Teori.....	7
2.2.1 <i>Stunting</i>	7
2.2.2 <i>Machine Learning</i>	8
2.2.3 Klasifikasi	10
2.2.4 <i>Support Vector Machine (SVM)</i>	11
2.2.5 <i>Feature Transformation</i>	13
2.2.6 Analisis Korelasi Fitur	14
2.2.7 <i>Cross Validation</i>	15
2.2.8 <i>Confusion Matrix</i>	18

III. METODOLOGI PENELITIAN.....	20
3.1 Tempat Dan Waktu Penelitian	20
3.1.1 Tempat	20
3.1.2 Waktu.....	20
3.2 Alat Dan Data.....	22
3.2.1 Alat	22
3.2.2 Data.....	24
3.3 Metode Penelitian	25
IV. HASIL DAN PEMBAHASAN.....	28
4.1 <i>Feature Transformation</i>	28
4.2 Analisis Korelasi Fitur	29
4.3 Percobaan Model Klasifikasi.....	31
4.3.1 <i>Import</i> Data dan Penganalisan Data.....	31
4.3.2 Pengelompokan dan Pemanggilan Data Berdasarkan Kelas.....	32
4.3.3 Pembagian Data Set	32
4.3.4 Membuat Model Klasifikasi.....	33
4.3.5 Klasifikasi Dengan SVM Menggunakan <i>Kernel Linear, Gaussian,</i> <i>dan Polynomial</i>	34
4.4 Pemodelan Hasil Prediksi	40
4.4.1 Tampilan Awal Sistem.....	40
4.4.2 Tampilan Hasil Analisis.....	41
4.5 Pembahasan.....	41
V. SIMPULAN DAN SARAN	44
5.1 Simpulan	44
5.2 Saran	45
DAFTAR PUSTAKA	46

DAFTAR GAMBAR

Gambar		Halaman
1.	Rata-rata Prevelensi Balita <i>Stunting</i> di Regional Asia Tenggara Tahun 2005-2017.	2
2.	Teknik Pembelajaran Pada <i>Machine Learning</i>	9
3.	Pendekatan Konseptual Pada Metode <i>Support Vector Machine</i>	12
4.	Contoh <i>K-Fold Cross Validation</i>	16
5.	Contoh <i>Hold-Out Cross Validation</i>	17
6.	Contoh <i>Leave One-Out Cross Validation</i>	17
7.	<i>Flowchart</i> Klasifikasi dengan Model SVM.	25
8.	Hasil Analisis Korelasi Fitur.	30
9.	Tampilan Awal Sistem.	40
10.	Tampilan Hasil Analisis.	41
11.	Grafik Hasil Prediksi Menggunakan <i>Kernel Linear, Gaussian</i> , dan <i>Polynomial</i>	42

DAFTAR TABEL

Tabel		Halaman
1.	Interpretasi Terhadap Koefisien Korelasi	15
2.	Tabel <i>Confusion Matrix</i>	18
3.	Alokasi Waktu Pelaksanaan Penelitian.....	21
4.	Contoh Data Penelitian	24
5.	Contoh Data Hasil <i>Feature Transformation</i>	29
6.	Hasil <i>Confusion Matrix</i> Menggunakan <i>Kernel Linear</i> Dengan <i>10-Fold Cross Validation</i>	35
7.	Hasil Prediksi Model SVM Klasifikasi Menggunakan <i>Kernel Linear</i> Dengan <i>10-Fold Cross Validation</i>	36
8.	Hasil <i>Confusion Matrix</i> Menggunakan <i>Kernel Gaussian</i> Dengan <i>10-Fold Cross Validation</i>	37
9.	Hasil Prediksi Model SVM Klasifikasi Menggunakan <i>Kernel Gaussian</i> Dengan <i>10-Fold Cross Validation</i>	37
10.	Hasil <i>Confusion Matrix</i> Menggunakan <i>Kernel Polynomial</i> Dengan <i>10-Fold Cross Validation</i>	38
11.	Hasil Prediksi Model SVM Klasifikasi Menggunakan <i>Kernel Polynomial</i> Dengan <i>10-Fold Cross Validation</i>	39

DAFTAR KODE PROGRAM

Kode Program	Halaman
1. Kode Menampilkan Hasil Analisis Korelasi Fitur.	29
2. Kode <i>Import</i> Data dan Penganalisan Data.....	31
3. Kode Pengelompokan dan Pemanggilan Data Berdasarkan Kelas. ...	32
4. Kode Pembagian Dataset.	33
5. Kode Membuat Model Klasifikasi.	33
6. Kode Melakukan Prediksi dengan Model Klasifikasi.	34
7. Kode Mencari Kinerja Dari Hasil Prediksi.	34

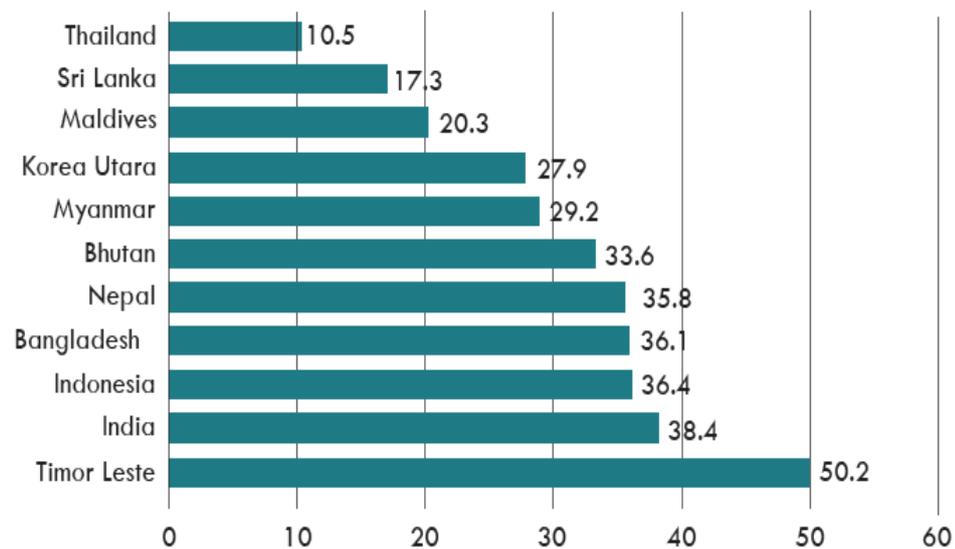
I. PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan (*growth*) merupakan masalah perubahan dalam ukuran besar, jumlah, dimensi tingkat sel, organ, dan individu yang dapat diukur dengan ukuran berat (gram, kilogram) dan ukuran panjang (centimeter, meter). Pertumbuhan menjadi dampak terhadap aspek fisik seseorang (Riyadi & Sukarmin, 2013). Salah satu dampak aspek fisik pada sistem pertumbuhan adalah status gizi pendek yang disebut dengan *stunting*. Kondisi ini terjadi apabila tinggi badan seseorang tidak sesuai dengan usia yang dimilikinya (Sakti, 2018). *Stunting* memiliki proporsi usia dengan rentang 48-59 bulan. Hal ini menunjukkan bahwa *stunting* rata-rata menyerang bayi di bawah lima tahun (balita). Semakin tinggi usia anak, kebutuhan nutrisi dan energi yang didapat juga harus semakin meningkat. Persediaan dan pemberian nutrisi yang cukup harus dilakukan agar pertumbuhan pada anak tidak menyimpang dari kondisi normal (Hanum, Khomsan, dan Heryatno, 2014).

Saat ini, *stunting* menjadi salah satu permasalahan yang akan diselesaikan oleh pemerintah Indonesia pada periode 2019-2024. Pemerintah Indonesia ingin menurunkan angka *stunting* agar anak-anak Indonesia dapat tumbuh menjadi generasi yang bermutu (Antons, 2019). Pusat data dan informasi Kementerian Kesehatan Republik Indonesia mencantumkan hasil riset mengenai *stunting* yang terbagi menjadi dua situasi, yaitu situasi global dan situasi nasional. Menurut hasil riset World Health Organization (WHO) pada tahun 2005-2017, lebih dari setengah balita penderita *stunting* di dunia berasal dari Asia dengan rata-rata prevalensi (jumlah individu dalam populasi yang mengalami penyakit, gangguan, atau kondisi tertentu) sebesar

55% atau sebanyak 83,6 juta balita. Data tersebut juga menjelaskan balita penderita *stunting* di regional asia memiliki rata-rata prevelensi terbesar pada negara Timor Leste dengan presentasi 50,2%, dan posisi kedua terbesar pada negara India dengan presentasi 38,4%, sedangkan negara Indonesia pada urutan ketiga terbesar dengan presentasi 36,4% (Sakti, 2018). Rata-rata prevelensi balita *stunting* di regional Asia Tenggara dapat dilihat pada Gambar 1.



Gambar 1. Rata-rata Prevelensi Balita *Stunting* di Regional Asia Tenggara Tahun 2005-2017 (Sakti, 2018).

Pada situasi nasional menurut survei Pemantauan Status Gizi (PSG) pada tahun 2017 di Indonesia, prevelensi anak-anak dengan tinggi badan di bawah rata-rata memiliki presentasi sebesar 29,6% (Sakti, 2018). Berdasarkan data tersebut, Indonesia memiliki jumlah penderita penyakit *stunting* yang cukup tinggi. Tingginya angka penderita *stunting* di Indonesia disebabkan oleh dua faktor, yaitu faktor luar dan faktor dalam. Faktor luar seperti kondisi lingkungan saat kehamilan, dan faktor dalam seperti konsumsi gizi yang diterima, dan faktor genetik yang diturunkan atau yang berasal dari orang tua (Astari, Nasoetion, dan Dwiriani, 2006).

Saat ini, diagnosis dan pemantauan terkait *stunting* masih banyak dilakukan sebatas mengukur langsung indikator gizi seperti dengan penanda biokimia dari beberapa nutrisi. Proses tersebut memang sangat penting dilakukan untuk mendapatkan hasil yang sesuai, tetapi sampai batas tertentu dan memerlukan biaya yang tidak sedikit (Hidayat, 2013). Permasalahan inilah yang menjadi fokus pada penelitian ini. Kemajuan teknologi dalam komputasi dan program pengolahan data memberikan kesempatan menggunakan data *non-nutritional*. Data *non-nutritional* meliputi umur, jenis kelamin, berat badan, panjang badan, dan lain-lain. Data *non-nutritional* yang tersedia ini dapat digunakan sebagai bahan pendukung penelitian dalam melakukan proses klasifikasi terkait penderita *stunting*. Salah satu penelitian yang pernah dilakukan terkait penyakit *stunting* adalah “*Backward Elimination* untuk Meningkatkan Akurasi Kejadian *Stunting* dengan Analisis Algoritma *Support Vector Machine*”. Penelitian ini bertujuan untuk memberikan keilmuan dalam pengelolaan data dalam mengambil suatu keputusan, serta untuk mengurangi dampak tentang kejadian *stunting*. Penelitian ini juga menghasilkan tingkat akurasi sebesar 81,62% dan nilai AUC (*Area Under Curve*) sebesar 0,921 dengan tingkat diagnose *Excellent Classification*. Selain itu, penambahan metode *backward elimination* meningkatkan nilai akurasi sebesar 90,16% dan nilai AUC sebesar 0,962 dengan tingkat diagnosa *Excellent Classification* (Byna dan Anisa, 2018).

Menurut penelitian Hears (1998) penggunaan metode *support vector machine* (SVM) dilakukan secara analitis (analisis secara matematis) dan memiliki kelebihan yaitu kemampuan generalisasi yang baik. Kemampuan generalisasi digunakan untuk dapat mengklasifikasikan suatu *pattern* yang tidak termasuk data yang digunakan pada fase pembelajaran metode tersebut. Oleh karena itu, penelitian ini menggunakan *support vector machine* sebagai metode dalam melakukan proses klasifikasi mengenai penderita *stunting* dengan menggunakan data lima puskesmas Kota Bandar Lampung.

1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang, maka rumusan masalah dari penelitian ini adalah bagaimana hasil kinerja dari metode *support vector machine* (SVM) dalam mengklasifikasikan penderita *stunting* menggunakan data lima puskesmas di Kota Bandar Lampung.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data sampel yang digunakan bersumber dari lima puskesmas di Kota Bandar Lampung tahun 2019.
2. Penentuan dalam melakukan proses klasifikasi penderita *stunting* menggunakan metode *support vector machine* (SVM) dengan 3 *kernel* yaitu, *kernel linear*, *kernel gaussian*, dan *kernel polynomial*.
3. Mengevaluasi hasil kinerja metode *support vector machine* (SVM) menggunakan 3 parameter yaitu, *accuracy*, *recall*, *precision*, dan *F1 score*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memberikan model prediksi untuk klasifikasi penderita *stunting* menggunakan data lima puskesmas di Kota Bandar Lampung.
2. Menentukan dan mengukur perbandingan dari hasil analisis korelasi fitur terhadap variabel.
3. Mengukur dan mengevaluasi hasil kinerja metode *support vector machine* (SVM) dalam melakukan proses klasifikasi penderita *stunting* berupa *accuracy*, *recall*, *precision*, dan *F1 score*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mengetahui hasil dan evaluasi dari kinerja proses klasifikasi penderita *stunting* menggunakan metode *support vector machine* (SVM).
2. Memberikan model yang dapat membantu pemerintah dan masyarakat dalam upaya guna mengantisipasi kemungkinan terjadinya *stunting*.

II. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Melakukan suatu penelitian diperlukan tambahan mengenai hasil-hasil penelitian yang telah ada sebelumnya, yang berkaitan dengan topik penelitian yang akan dilakukan dan dapat dibandingkan hasil dari penelitian tersebut. Berikut adalah sumber penelitian terdahulu yang telah dilakukan dan digunakan sebagai pendukung penelitian ini.

2.1.1 Penelitian oleh Hendra & Gazali (2016) meneliti tentang klasifikasi nutrisi balita di Kota Palu menggunakan *support vector machine* (SVM). Penelitian ini bertujuan untuk membantu tenaga kesehatan mengetahui status gizi anak secara terkomputerisasi dan membantu pemerintah dalam pengambilan kebijakan tentang gizi balita di Kota Palu. Penelitian ini menggunakan metode *support vector machine* (SVM) dengan data penelitian sebanyak 640 sampel data balita dari delapan puskesmas di Kota Palu. Hasil penelitian ini menunjukkan bahwa dari total 320 sampel data balita terdapat 295 data dari sistem yang sesuai dengan kondisi nyata, dan sisanya 25 data tidak sesuai. Penelitian ini juga menunjukkan tingkat akurasi yang didapat oleh sistem sebesar 92,18%.

2.1.2 Penelitian lainnya oleh Kusumaningrum dkk. (2020) meneliti tentang tolak ukur dari algoritma *multi-class* terhadap pengklasifikasian terkait penyakit *stunting*. Penelitian ini bertujuan mengevaluasi hasil

kinerja beberapa algoritma, untuk memberikan tolak ukur algoritma dalam pengolahan klasifikasi *multi-class*. Penelitian ini menggunakan sebanyak 900 sampel data untuk tiga kelas dengan parameter untuk setiap algoritmanya adalah *regression*, *random forest*, dan *support vector machine* (SVM). Hasil dari penelitian ini menunjukkan bahwa metode *support vector machine* mampu memberika hasil yang lebih tinggi dibandingkan dari kedua algoritma lain yang digunakan dalam pengolahan klasifikasi *multi-class*. Metode *support vector machine* pada penelitian ini menghasilkan nilai akurasi sebesar 98% dengan standar deviasi sebesar 0.03.

2.2 Uraian Landasan Teori

2.2.1 *Stunting*

Stunting merupakan salah satu permasalahan gizi yang tengah dihadapi di dunia, terkhusus pada negara-negara berkembang yang masih memiliki tingkat kemiskinan yang tinggi. Menjadi permasalahan yang cukup serius, *stunting* berhubungan dengan meningkatnya resiko terjadinya kesakitan dan kematian, dikarenakan perkembangan motorik terlambat dan terhambatnya pertumbuhan mental. *Stunting* atau Kegagalan tumbuh (*growth faltering*) terjadi akibat akumulasi ketidakcukupan nutrisi yang didapat selama masa awal dimulai dari proses kehamilan sampai usia 24 bulan (Mitra, 2015).

Kondisi kesehatan dan gizi yang dimiliki oleh ibu pada saat kehamilan serta setelah melakukan persalinan dapat mempengaruhi pertumbuhan janin dan resiko terjadinya *stunting*. Memiliki usia kehamilan yang masih terlalu muda dengan batas di bawah 20 tahun memiliki resiko melahirkan bayi dengan berat lahir rendah (BBLR). Bayi dengan

BBLR mempengaruhi sekitar 20% dari proses bayi terdampak *stunting*. Dari data Riskesdas pada tahun 2013, didapat bahwa proporsi kehamilan pada usia remaja (usia 10-14 tahun) sebanyak 0,02% dan (usia 15-19 tahun) sebanyak 1,97%. (Sakti, 2018).

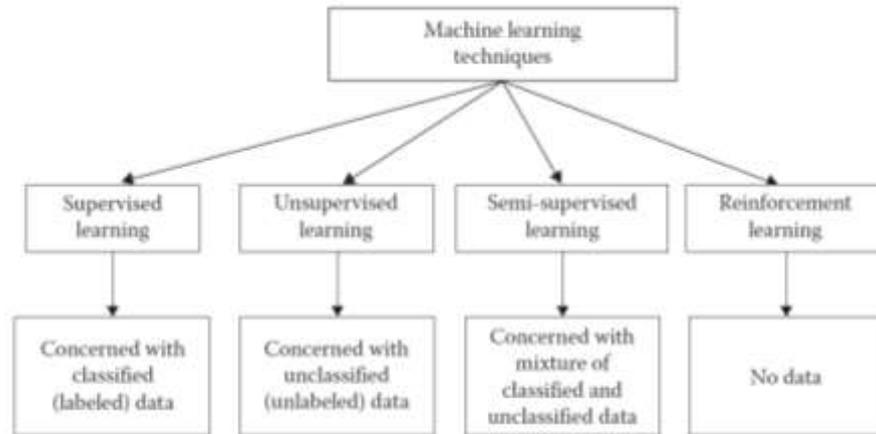
Penelitian ini menggunakan penyakit *stunting* sebagai salah satu objek yang akan diteliti, dan menggunakan variabel pendukung yang menentukan penyebab terjadinya penyakit *stunting*.

2.2.2 *Machine Learning*

Machine learning atau dalam bahasa Indonesia disebut dengan pembelajaran mesin merupakan cabang *artificial intelligence* yang memiliki tujuan untuk memungkinkan mesin melakukan pekerjaannya secara terampil dengan menggunakan perangkat lunak yang cerdas. Dalam penggunaannya *machine learning* memiliki algoritma yang memerlukan data untuk dipelajari, dengan begitu ilmu ini memerlukan hubungan dengan disiplin ilmu basis data, dalam hal *data mining*, dan juga pengenalan pola (Mohammed, Khan, dan Bashie, 2016).

Machine learning berguna dalam mengkarakterisasi masalah pembelajaran sesuai dengan jenis data yang digunakan. Hal ini sangat membantu dalam masalah dengan tipe data yang serupa dengan teknik yang sama, contoh pada pengguna pemrosesan bahasa alami dalam melakukan string teks bahasa alami dan penggunaan bioinformatika dalam melakukan sekuens DNA (Smola dan Vishwanathan, 2009).

Menurut penelitian Mohammed, Khan, dan Bashie (2016) *machine learning* terbagi menjadi empat teknik pembelajaran, seperti yang dapat dilihat pada Gambar 2.



Gambar 2. Teknik Pembelajaran Pada *Machine Learning* (Mohammed, Khan, dan Bashie, 2016).

Adapun penjelasan teknik pembelajaran pada *machine learning* (Mohammed, Khan, dan Bashie, 2016) seperti yang ditampilkan oleh Gambar 2, adalah sebagai berikut:

- *Supervised Learning*

Teknik pembelajaran *supervised learning* dalam menyimpulkan suatu fungsi atau pemetaan dari data yang akan diteliti menggunakan sebuah label. Data tersebut terdiri dari vektor input berupa X dan vektor keluaran berupa Y yang diberikan label. Dalam penggunaannya seringkali penilaian yang dilakukan manual oleh manusia memiliki tingkat akurasi yang lebih tinggi, namun dalam beberapa kasus, mesin juga dapat digunakan untuk pelabelan yang andal. Kategori algoritma dalam *supervised learning* diantaranya adalah *regression* dan *classification*.

- *Unsupervised Learning*

Berbeda dengan *supervised learning*, pada teknik pembelajaran *unsupervised learning* data yang akan diteliti tidak menggunakan label. Tujuan digunakan *unsupervised learning* adalah untuk menemukan struktur tersembunyi dari data tersebut. Kategori algoritma dalam *unsupervised learning* yaitu *clustering*.

- *Semi-Supervised Learning*

Teknik pembelajaran *semi-supervised learning*, data yang diberikan adalah data campuran, berupa data *classified* yang tidak terklasifikasi. Kombinasi ini digunakan untuk menghasilkan suatu model yang sesuai untuk klasifikasi data. Disebagian besar situasi, data berlabel adalah data yang sulit didapatkan dan data yang tidak berlabel adalah data yang cukup mudah ditemukan.

- *Reinforcement Learning*

Teknik pembelajaran *reinforcement learning* bertujuan dalam menggunakan data observasi yang telah dikumpulkan, dari interaksi dengan lingkungan untuk mengambil tindakan yang digunakan dalam meminimalkan resiko dan menghasilkan program yang cerdas.

2.2.3 Klasifikasi

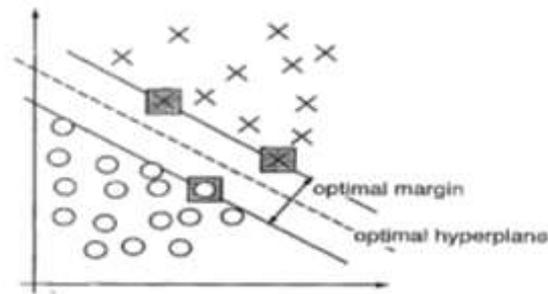
Klasifikasi adalah suatu sistem pengelompokan atau penggolongan sejumlah objek, gagasan, buku atau benda-benda lain berdasarkan subjek atau dengan melihat ciri-ciri yang sama. Klasifikasi berguna agar dalam penyusunannya dapat dilakukan secara teratur dengan kesamaan subjek yang saling berdekatan letaknya, dan memisahkan subjek yang memiliki ciri-ciri berbeda (Saputro, 2017).

Klasifikasi memiliki tujuan untuk dapat mengorganisasikan suatu objek dengan sistem tertentu sehingga mudah diketemukan. Adapun tujuan klasifikasi secara rinci menurut Subrata (2009) adalah sebagai berikut:

- Menghasilkan Urutan Yang Berguna
Tujuan utama klasifikasi adalah menghasilkan urutan atau susunan yang berguna agar memudahkan saat dilakukan identifikasi.
- Penempatan Yang Tepat
Pada bahan pustaka apabila diperlukan pemakaian kembali, objek lebih mudah diketemukan serta mudah dikembalikan ke bentuk semula sesuai dengan sistem klasifikasi yang digunakan.
- Penyusunan Mekanis
Objek klasifikasi apabila dilakukan penarikan ataupun penyisipan tidak akan mengganggu susunan yang telah dibuat.

2.2.4 *Support Vector Machine (SVM)*

Support vector machine adalah salah satu metode yang banyak digunakan dalam menyelesaikan masalah klasifikasi untuk berbagai bidang. Metode *support vector machine* memiliki dua pendekatan yaitu secara konseptual dan secara teknis. Secara konseptual dengan menentukan suatu *hyperplane* (sebuah garis lurus atau bidang mendatar yang memisahkan suatu kelas atau sebuah batas keputusan) yang memaksimalkan margin antara dua kelas data. Secara teknis dengan mentransformasikan data ke ruang dimensi yang lebih tinggi dengan menggunakan *kernel* (Cortes dan Vapnik, 1995). Konseptual mengenai metode SVM dapat dilihat pada Gambar 3.



Gambar 3. Pendekatan Konseptual Pada Metode *Support Vector Machine* (Cortes dan Vapnik, 1995).

Kinerja dari algoritma metode *support vector machine* tergantung pada pemilihan variabel yang digunakan. Algoritma *support vector machine* bertujuan untuk menemukan fungsi linear dari bentuk $f(x) = w^T x + b$, dengan $w \in R^p$ dan $b \in R$ sehingga titik data x yang telah ditetapkan akan memberikan nilai label +1 jika $f(x) > 0$, dan akan memberikan nilai label -1 jika $f(x) < 0$ (Dong, Xia, dan Tu, 2007).

Menurut penelitian Liu, Shen, & Wang (2014) pada *support vector machine* terdapat empat *kernel* dasar yang dapat digunakan, sebagaimana persamaan 1, 2, 3, dan 4:

- *Linear kernel function:*

$$K(X_i, X_j) = X_i \cdot X_j \dots\dots\dots(1)$$

- *Polynomial kernel function:*

$$K(X_i, X_j) = (X_i \cdot X_j + 1)^h \dots\dots\dots(2)$$

- *Gaussian radial basis function kernel function:*

$$K(X_i, X_j) = e^{-\frac{\|X_i - X_j\|^2}{2\sigma^2}} \dots\dots\dots(3)$$

- *Sigmoid kernel function:*

$$K(X_i, X_j) = \tanh(\kappa X_i \cdot X_j - \delta) \dots\dots\dots(4)$$

Penelitian ini menggunakan metode *support vector machine* dalam mengklasifikasikan penderita *stunting* dan menggunakan *kernel linear, polynomial, dan gaussian*.

2.2.5 Feature Transformation

Feature transformation adalah proses membuat set variabel baru dengan ukuran yang lebih kecil. Tujuan dari *feature transformation* untuk mengambil informasi yang paling berguna, yang terdapat pada set variabel asli saat memprediksi hasil. Set variabel baru yang didapatkan merupakan hasil penerapan *transformation* dari set variabel asli. Variabel yang ditransformasikan merupakan perwakilan dari proyeksi set variabel asli yang dirubah ke ruang set variabel baru, dengan kelompok hasil yang berbeda dengan pemisahan yang lebih baik dibandingkan dengan ruang set variabel asli (Awan dkk., 2018).

Proses pada *feature transformation* sangat penting untuk menentukan jenis variabel yang ada pada setiap fitur dengan variabel hasil *transformation* berupa variabel numerik. Menurut penelitian Vickery (2018) variabel numerik dalam ilmu statistika, dikarakterisasi menjadi empat jenis data, sebagai berikut:

- *Continuous*
Variabel *continuous* adalah variabel yang memiliki jumlah kemungkinan nilai yang tidak terbatas atau tidak memiliki rentang nilai tertentu.
- Nominal
Variabel nominal adalah variabel yang memiliki nilai kategorikal dengan dua kemungkinan nilai atau lebih, tetapi urutan nilai tersebut tidak memiliki arti.

- *Dichotomous*
Variabel *dichotomous* adalah variabel yang memiliki nilai kategorikal, akan tetapi hanya memiliki dua kemungkinan nilai. Kemungkinan nilai tersebut biasanya berupa nilai 0 dan 1.
- *Ordinal*
Variabel ordinal adalah variabel yang memiliki 2 atau lebih kemungkinan nilai, akan tetapi nilai-nilai tersebut memiliki urutan atau peringkat yang berarti.

2.2.6 Analisis Korelasi Fitur

Analisis korelasi fitur adalah adalah suatu metode yang digunakan dalam menentukan besaran suatu hubungan antara variabel satu dengan variabel lainnya, tanpa mempersoalkan apakah variabel tersebut saling bergantung. Terdapat dua beberapa teknik korelasi yang sangat sering digunakan yaitu korelasi pearson dan korelasi spearman. Pada analisis korelasi fitur terdapat suatu ukuran yang digunakan untuk mengetahui derajat hubungan antar variabel yaitu disebut dengan koefisien korelasi. Nilai koefisien korelasi berada diantara $-1 < 0 < 1$ yaitu apabila $r = -1$ didapatkan nilai korelasi negatif sempurna yang artinya korelasi tersebut memiliki hubungan yang berlawanan, dan apabila $r = 1$ didapatkan korelasi positif yang artinya korelasi tersebut memiliki hubungan yang searah. Jika koefisien korelasi bernilai 0, maka tidak didapatkan hubungan antar variabel (Safitri, 2016).

Menurut penelitian Iin dan Sugiarto (2012) dipaparkan suatu interpretasi terhadap koefisien korelasi yang dapat digunakan sebagai acuan untuk mengetahui tinggi atau lemahnya hubungan korelasi antar variabel, yang dapat dilihat pada tabel 1.

Tabel 1. Interpretasi Terhadap Koefisien Korelasi (Iin & Sugiarto, 2012)

Besar Koefisien Korelasi (Positif atau Negatif)	Interpretasi Koefisien Korelasi
0,00	Tidak Ada Korelasi
0,01 – 0,20	Korelasi Sangat Lemah
0,21 – 0,40	Korelasi Lemah
0,41 – 0,70	Korelasi Sedang
0,71 – 0,99	Korelasi Tinggi
1,00	Korelasi Sempurna

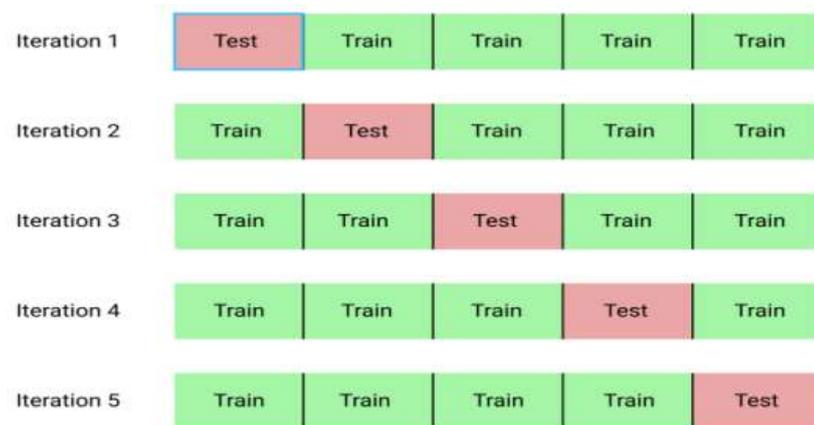
2.2.7 Cross Validation

Cross validation adalah sebuah metode statistika yang digunakan untuk mengevaluasi dan membandingkan algoritma pembelajaran dengan membaginya menjadi dua bagian yaitu untuk melakukan pembelajaran dan pelatihan untuk model yang lain, dan juga digunakan dalam memvalidasi model (Refaeilzadeh et al., 2008). Berikut adalah jenis-jenis dari *cross validation*.

- *K-Fold Cross Validation*

K-fold cross validation adalah teknik umum untuk memperkirakan kinerja pengklasifikasian. Proses *k-fold cross validation* dilakukan dengan, data pertama kali akan dipartisi menjadi k yang sama dengan ukuran iterasi atau lipatan. Selanjutnya k dalam model pelatihan dan validasi dilakukan pengulangan (iterasi) sedemikian rupa sehingga dalam setiap iterasi, iterasi data yang berbeda tidak berlaku untuk dilakukan validasi. Dalam penggunaan jumlah *fold* terbaik pada *k-fold cross validation* untuk dilakukan uji validitas, dianjurkan menggunakan 10 *k-fold cross validation* ($k = 10$) dalam model tersebut

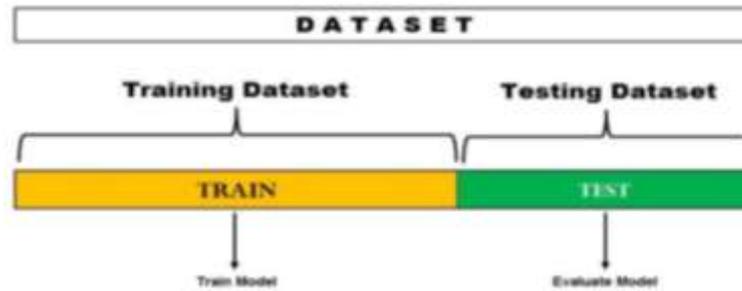
(Refaeilzadeh et al., 2008). Berikut contoh *k-fold cross validation* yang dapat dilihat pada Gambar 4.



Gambar 4. Contoh *K-Fold Cross Validation* (Azis dkk., 2020).

- *Hold-Out Cross Validation*

Hold-out cross validation adalah teknik validasi yang paling sederhana, pada metode ini data dibagi secara acak menjadi dua bagian. Bagian ini disebut data pelatihan (*training*) dan data tes (*validation*), dengan rasio tipikal pembagian 60:40 atau 80:20. Teknik ini digunakan untuk melatih model menggunakan data pelatihan dan mengevaluasi model menggunakan data tes. Teknik evaluasi model yang digunakan pada data pelatihan dalam menghitung kesalahan bergantung dari jenis masalah pada penelitian, jika data pelatihan atau pengujian tidak dapat merepresentasikan data secara lengkap maka hasil data dapat menjadi tidak baik (Yadav & Shukla, 2016). Berikut contoh *hold-out cross validation* dapat dilihat pada Gambar 5.



Gambar 5. Contoh *Hold-Out Cross Validation* (Yadav & Shukla, 2016).

- *Leave-One-Out Cross Validation*

Leave-one-out cross validation adalah special kasus *k-fold cross validation* dimana k sama dengan jumlah *instance* dalam data. Dengan kata lain disetiap iterasi hampir semua data kecuali untuk observasi data tunggal, data digunakan untuk pelatihan modelnya dan model yang digunakan diuji dengan pengamatan tunggal tersebut. Metode ini masih banyak digunakan ketika tersedia data yang sangat jarang, terutama pada bidang bioinformatika yang penerapannya masih menggunakan sampel data yang sedikit tersedia (Refaeilzadeh et al., 2008). Berikut contoh *leave-one-out cross validation* yang dapat dilihat pada Gambar 6.



Gambar 6. Contoh *Leave One-Out Cross Validation* (Refaeilzadeh et al., 2008).

2.2.8 Confusion Matrix

Confusion matrix berisi informasi mengenai hasil klasifikasi aktual dan hasil klasifikasi prediksi yang dilakukan oleh sistem. *Confusion matrix* digunakan untuk mengevaluasi kinerja suatu sistem dengan menyajikan data dalam matriks (Santra & Christy, 2012). *Confusion matrix* dapat dilihat pada Tabel 2.

Tabel 2. Tabel *Confusion Matrix* (Kohavi, R. & Provost, 1998)

		<i>Actual Class</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Predicted Class</i>	<i>Positive</i>	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	<i>Negative</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Menurut penelitian Kohavi, R. & Provost (1998) Tabel 2 dapat dijelaskan sebagai berikut:

- *True Positive* adalah apabila memprediksi *positive* dan hasil tersebut benar.
- *False Positive* adalah apabila memprediksi *positive* dan hasil tersebut salah.
- *False Negative* adalah apabila memprediksi *negative* dan hasil tersebut salah.
- *True Negative* adalah apabila memprediksi *negative* dan hasil tersebut benar.

Menurut penelitian Kohavi, R. & Provost (1998) juga dijelaskan bahwa nilai yang dihasilkan melalui metode *confusion matrix* adalah berupa evaluasi sebagai berikut:

- *Accuracy*

Accuracy adalah presentase seberapa akuratnya suatu model dalam mengklasifikasikan (prediksi) secara benar oleh algoritma.

Fungsi dari *accuracy* dapat dilihat pada persamaan 5:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \times 100\% \dots\dots\dots (5)$$

- *Recall* atau *Sensitivity*

Recall adalah presentase keberhasilan suatu model dalam menemukan kembali (observasi) sebuah informasi. Fungsi dari *recall* dapat dilihat pada persamaan 6:

$$Recall = \frac{TP}{(TP+FN)} \times 100\% \dots\dots\dots (6)$$

- *Precision*

Precision adalah presentase tingkat ketepatan antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Fungsi dari *precision* dapat dilihat pada persamaan 7:

$$Precision = \frac{TP}{(TP+FP)} \times 100\% \dots\dots\dots (7)$$

- *F1 Score*

F1 score adalah perbandingan rata-rata nilai *recall* dan *precision* yang dibobotkan. *F1 score* digunakan apabila ingin membuat model klasifikasi dengan keseimbangan *recall* dan *precision* secara optimal. Fungsi dari *F1 score* dapat dilihat pada persamaan 8:

$$F1\ Score = \frac{2 \times recall \times precision}{(recall+precision)} \times 100\% \dots\dots\dots (8)$$

III. METODOLOGI PENELITIAN

3.1 Tempat Dan Waktu Penelitian

3.1.1 Tempat

3.1.1.1 Tempat Pengumpulan Data

Proses pengumpulan data pada penelitian ini diperoleh dari lima puskesmas yang tersebar di Kota Bandar Lampung, yaitu Puskesmas Way Halim, Puskesmas Sukaraja, Puskesmas Kota Karang, Puskesmas Pasar Ambon, dan Puskesmas Kemiling.

3.1.1.2 Tempat Penelitian

Penelitian ini juga dilakukan di Laboratorium Rekayasa Perangkat Lunak (RPL) Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang beralamatkan di jalan Soemantri Brojonegoro No. 1 Gedung Meneng, Bandar Lampung.

3.1.2 Waktu

Penelitian ini dilaksanakan mulai bulan Februari 2020 sampai dengan bulan April 2021. Rincian alokasi waktu penelitian dapat dilihat pada Tabel 3.

Perencanaan waktu yang digunakan pada penelitian ini dimulai pada bulan Februari. Proses awal berupa pengumpulan data yang diperoleh dari lima puskesmas yang tersebar di Kota Bandar Lampung. Setelah data penelitian diperoleh, data tersebut akan dilakukan proses *cleaning* pada bulan April sampai bulan Juni untuk menghapus data kosong serta redundansi data. Proses selanjutnya yaitu *feature transformation* yang dilakukan pada bulan Juni sampai bulan Juli. Proses ini digunakan untuk mengubah data nominal menjadi data numerik, supaya data tersebut memiliki nilai. Pada bulan Agustus sampai Bulan September, proses *k-fold cross validation* dilakukan untuk uji validitas dengan *10-fold cross validation* yang membagi data secara acak menjadi 10 iterasi. Proses kelima pada bulan Oktober sampai bulan November yaitu membangun model klasifikasi guna mendapat nilai akurasi dan hasil prediksi. Proses terakhir pada pada bulan Desember sampai bulan Januari yaitu melakukan proses *testing* atau pengujian menggunakan *confusion matrix* untuk mendapatkan hasil terbaik dari pemodelan menggunakan metode SVM.

Penelitian ini juga memuat alokasi waktu untuk studi literatur & penulisan laporan bab 1-3 pada bulan Februari sampai bulan Juni, dengan seminar usul penelitian pada bulan Juni. Penulisan laporan bab 4-5 dilakukan pada bulan Desember sampai pada bulan April, dengan seminar hasil penelitian pada bulan April.

3.2 Alat Dan Data

3.2.1 Alat

3.2.1.1 Perangkat Keras

- a. Prosesor: *Intel®Core™ i3-3217U CPU @ 1.80GHz to 1.80GHz.*

- b. Memori: DDR3 1333 MHz SDRAM, OnBoard Memory 4 GB.
- c. *System Type: 64-bit operating system, x64-based Processor.*

3.2.1.2 Perangkat Lunak

- a. Sistem Operasi Windows 10 64-bit digunakan untuk mengatur dan mengontrol kerja dari perangkat keras dan menjalankan aplikasi pada sistem komputer.
- b. RStudio versi 3.5.1 digunakan sebagai sistem perangkat lunak untuk mengolah dan menganalisis data penelitian. Adapun *package* pada R Studio yang digunakan penelitian ini adalah sebagai berikut:
 - *Package readr* versi 1.4.0 (Wickham, 2020), digunakan untuk memasukkan data file.
 - *Package caret* versi 6.0-86 (Kuhn, 2020), digunakan untuk fungsi *confusion matrix*.
 - *Package e1071* versi 1.7-4 (Meyer, 2020), digunakan untuk analisis kelas pada *support vector machine*.
 - *Package psych* versi 2.0.12 (Revelle, 2020), digunakan untuk menampilkan grafik pada proses analisis korelasi fitur.
 - *Package shiny* versi 1.6.0 (Chang, 2021), digunakan untuk membuat aplikasi web yang interaktif dengan mudah.
- c. Microsoft Excel 2010 digunakan untuk memuat dan menampilkan data penelitian.

3.2.2 Data

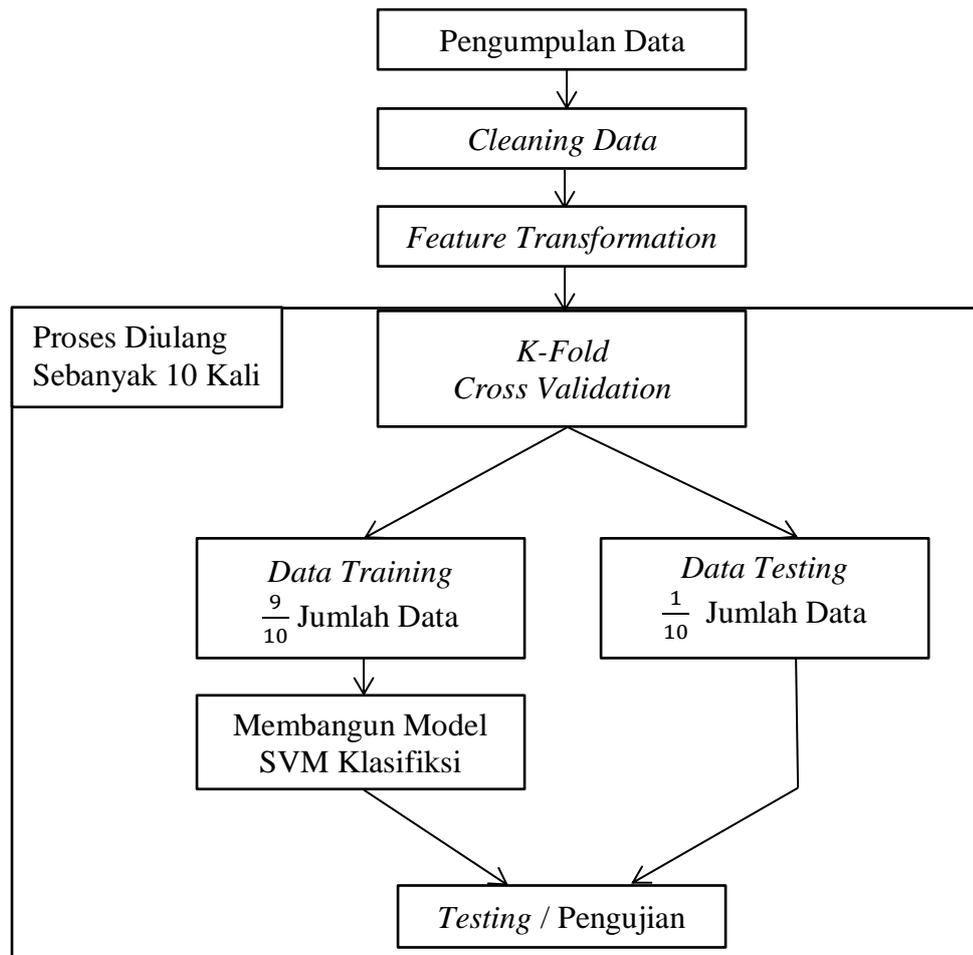
Data yang digunakan pada penelitian berupa 234 data sampel balita pada tahun 2019 yang terdiri dari 89 data dengan kejadian *stunting*, dan 145 data dengan kejadian tidak *stunting*. Data penelitian ini diperoleh dari 5 puskesmas di Kota Bandar Lampung, diantaranya 9 data dari Puskesmas Way Halim, 34 data dari Puskesmas Sukaraja, 70 data dari Puskesmas Kota Karang, 12 data dari Puskesmas Pasar Ambon, dan 109 data dari Puskesmas Kemiling. Data terdiri dari tujuh atribut yaitu, jenis kelamin, usia, berat badan (BB), panjang badan (PB), berat badan lahir (BBL), panjang badan lahir (PBL), dan pemberian asi, serta terdapat satu kelas yaitu *stunting*. Contoh dari data penelitian dapat dilihat pada Tabel 4.

Tabel 4. Contoh Data Penelitian

JK	Umur (Bulan)	BBL (Kg)	PBL (Cm)	BB (Kg)	PB (Cm)	Pemberian Asi	Stunting
L	18	3	49	9	76	Tidak	Ya
L	15	2,8	49	10	71	Tidak	Ya
L	11	3,1	50	9	69	Tidak	Ya
L	11	2,9	49	9	69	Tidak	Ya
P	8	3	49	8	63	Tidak	Ya
P	7	2,8	49	7,7	65	Tidak	Tidak
P	14	2,7	49	7,2	65	Tidak	Ya
L	12	2,9	49	2,9	49	Tidak	Ya
L	6	3	49	6,2	63	Tidak	Ya
P	48	3	49	10,5	92,7	Tidak	Ya
L	17	2,9	48	8,5	65	Tidak	Ya
P	16	2	49	7,1	72,2	Tidak	Ya
L	9	2,3	50	6,4	68,3	Tidak	Tidak
L	18	2,8	49	7	73,5	Tidak	Ya
L	33	2,8	49	8,4	80	Tidak	Ya

3.3 Metode Penelitian

Penelitian yang dilakukan untuk klasifikasi penderita *stunting* di Kota Bandar Lampung dengan metode *support vector machine*, dilakukan dengan beberapa tahap yang dapat dilihat melalui proses yang ditunjukkan pada Gambar 7.



Gambar 7. Flowchart Klasifikasi dengan Model SVM.

Berdasarkan *flowchart* klasifikasi pada Gambar 7, berikut merupakan penjelasan dari setiap proses yang akan dilakukan pada penelitian ini.

- Proses 1 – Pengumpulan Data
Proses pertama dalam melakukan penelitian ini adalah dengan melakukan pengumpulan data yang berkaitan dengan penelitian. Data

yang digunakan adalah 234 data sampel pada balita yang diperoleh dari lima puskesmas di Kota Bandar Lampung pada tahun 2019. Data terdiri dari tujuh atribut yaitu, jenis kelamin, usia, berat badan (BB), panjang badan (PB), berat badan lahir (BBL), panjang badan lahir (PBL), dan pemberian asi.

- Proses 2 – *Cleaning Data*

Proses kedua adalah *cleaning* data, setelah data dikumpulkan kemudian dilakukan pembersihan data dengan menghapus data yang tidak sesuai untuk dilakukan penelitian, seperti menghapus data kosong yang tidak memiliki nilai, menghilangkan redundansi data, dan mengambil data penting yang akan diteliti. Proses *cleaning* data pada penelitian ini dilakukan secara manual dengan menghapus sebanyak 331 data kosong dari total 565 data yang telah berhasil dikumpulkan.

- Proses 3 - *Feature Transformation*

Proses ketiga adalah dengan melakukan proses *feature transformation* yang dilakukan dengan mengubah data nominal menjadi data numerik. Pada penelitian ini proses, *feature transformation* dilakukan secara manual dengan memberi nilai pada data nominal seperti contoh pada atribut pemberian asi yang terbagi menjadi dua kategori Ya dan Tidak, dengan memberi nilai Ya = 1 dan Tidak = 0.

- Proses 4 – *K-Fold Cross Validation*

Proses selanjutnya adalah dengan membagi data penelitian menjadi data *testing* dan data *training*. Pembagian data yaitu, $\frac{1}{10}$ data sebagai data *testing*, dan $\frac{9}{10}$ data sebagai data *training*. Uji validitas menggunakan *10-fold cross validation* dengan membagi data secara acak menjadi 10 iterasi. Proses diulangi sampai setiap iterasi k yang berfungsi sebagai set uji dan mengambil rata-rata skor yang tercatat.

- Proses 5 – Membangun Model SVM Klasifikasi
Pada proses ini, hal yang terlebih dahulu dilakukan yaitu dengan menentukan tipe *kernel* dan nilai parameter. Selanjutnya mulai dilakukan proses pembelajaran untuk model *support vector machine* pada data *training*. Jika model terbaik pada data telah terpilih maka sudah bisa untuk dilakukan proses klasifikasi.
- Proses 6 – *Testing* / Pengujian
Proses keenam adalah melakukan *testing* atau pengujian. Model SVM yang telah didapatkan pada tahapan sebelumnya diuji menggunakan data *testing* dari proses pembagian data penelitian. Pengujian yang digunakan dengan menggunakan *confusion matrix* untuk mendapatkan nilai *accuracy*, *recall*, *precision*, dan *F1 score*.

V. SIMPULAN DAN SARAN

5.1 Simpulan

Dari penelitian yang dilakukan, didapatkan hasil dan dapat diambil simpulan sebagai berikut:

1. Model prediksi untuk klasifikasi penderita *stunting* menggunakan data lima puskesmas di Kota Bandar Lampung sudah berhasil dibuat.
2. Analisis korelasi fitur pada penelitian ini didapatkan hasil korelasi seluruhnya bernilai positif, yang artinya memiliki hubungan antara variabel secara teratur dengan arah yang sama. Hasil analisis korelasi tersebut juga didapatkan korelasi terbesar (korelasi tinggi) meliputi hubungan umur-panjang badan dengan nilai 0,92. Korelasi terkecil (korelasi sangat lemah) meliputi hubungan jenis kelamin-umur dan berat badan lahir-asi dengan nilai 0,04. Kemudian rata-rata dari hasil keseluruhan korelasi didapatkan nilai sebesar 0,23.
3. Hasil klasifikasi menggunakan metode *support vector machine* (SVM) didapatkan hasil rata-rata *accuracy* pada *kernel linear* sebesar 80,8%, hasil rata-rata *recall* sebesar 86,2%, hasil rata-rata *precision* sebesar 82,8%, dan hasil rata-rata *F1 score* sebesar 84,6%. Kemudian hasil rata-rata *accuracy* pada *kernel gaussian* sebesar 80,4%, hasil rata-rata *recall* sebesar 82%, hasil rata-rata *precision* sebesar 86%, dan hasil rata-rata *F1 score* sebesar 83,4%. Terakhir hasil rata-rata *accuracy* pada *kernel polynomial* sebesar 69,1%, dengan hasil rata-rata *recall* sebesar 90,3%, hasil rata-rata *precision* sebesar 69,6%, dan hasil rata-rata *F1 score* sebesar 78,3%.

5.2 Saran

Adapun saran yang dapat diberikan dalam penelitian ini adalah sebagai berikut:

1. Untuk penelitian selanjutnya agar dapat dilakukan penambahan jumlah data dan tahun penelitian guna mendapatkan nilai akurasi yang lebih baik.
2. Pada penelitian selanjutnya dapat menggunakan metode klasifikasi lainnya untuk dapat dibandingkan dengan hasil nilai akurasi pada penelitian ini.

DAFTAR PUSTAKA

- Antons. 2019. SDM Unggul Indonesia Maju. di antaranews.com website: <https://www.antaranews.com/infografik/1017556/sdm-unggul-indonesia-maju> (Akses September 13, 2019).
- Astari, L., Nasoetion, A., & Dwiriani, C. 2006. Hubungan Konsumsi Asi Dan Mp-Asi Serta Kejadian Stunting Anak Usia 6-12 Bulan Di Kabupaten Bogor. 30(1).
- Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., Chow, B. J., & Dwivedi, G. 2018. Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death. *PLoS ONE*, 14(6), 1–13. <https://doi.org/10.1371/journal.pone.0218760>.
- Azis, H., Purnawansyah, P., Fattah, F., & Putri, I. P. 2020. Performa Klasifikasi K-NN dan Cross Validation Pada Data Pasien Pengidap Penyakit Jantung. *ILKOM Jurnal Ilmiah*, 12(2), 81–86. <https://doi.org/10.33096/ilkom.v12i2.507.81-86>.
- Byna, A., & Anisa, F. N. 2018. Backward Elimination Untuk Meningkatkan Akurasi Kejadian Stunting Dengan Analisis Algoritma Support Vector Machine. 9(2), 55–63.
- Chang, W. 2021. shiny: Web Application Framework for R (1.6.0). website: <https://shiny.rstudio.com/> (Akses Juni 06, 2021).
- Cortes, C., & Vapnik, V. 1995. Support-Vector Networks. <https://doi.org/10.1109/64.163674>.
- Dong, Y., Xia, Z., & Tu, M. 2007. Selecting Optimal Parameters in Support Vector Machines. *Proceedings of the IEEE 6th International Conference on Machine Learning and Applications (ICMLA07)*.
- Hanum, F., Khomsan, A., & Heryatno, Y. 2014. Hubungan Asupan Gizi Dan Tinggi Badan Ibu Dengan Status Gizi Anak Balita. *Jurnal Gizi Dan Pangan*, 9(1), 1–6. <https://doi.org/10.25182/jgp.2014.9.1>.
- Hears, M. A. 1998. Support Vector Machines - Trends and controversies.

- Hendra, A., & Gazali, G. 2016. Support Vector Machine (SVM) For Toddler's Nutritional Classification in Palu City. *Insist*, 1(1), 49. <https://doi.org/10.23960/ins.v1i1.19>.
- Hidayat, R. 2013. Sistem Prediksi Status Gizi Balita Dengan Menggunakan Support Vector Regression. Departemen Ilmu Komputer Fakultas Matematika Dan Ilmu Pengetahuan Alam Institut Pertanian Bogor.
- In, Y., & Sugiarto, B. 2012. Korelasi Antara Keterampilan Metakognitif Dengan Hasil Belajar Siswa Di Sman 1 Dawarblandong, Mojokerto. *Unesa Journal of Chemical Education*, 1, 78–83.
- Kohavi, R. & Provost, F. 1998. Glossary of terms: Machine Learning. *Machine Learning*, 30, 271–274.
- Kuhn, M. 2020. caret: Classification and Regression Training (6.0-86). website: <https://github.com/topepo/caret/> (Akses Juni 06, 2021).
- Kusumaningrum, R., Indihatmoko, T. A., Juwita, S. R., Hanifah, A. F., Khadijah, K., & Surarso, B. 2020. Benchmarking of multi-class algorithms for classifying documents related to stunting. *Applied Sciences (Switzerland)*, 10(23), 1–13. <https://doi.org/10.3390/app10238621>.
- Liu, L., Shen, B., & Wang, X. 2014. Research on kernel function of support vector machine. *Lecture Notes in Electrical Engineering*, 260 LNEE, 827–834. https://doi.org/10.1007/978-94-007-7262-5_93.
- Meyer, D. 2020. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien (1.7-4).
- Mitra. 2015. Permasalahan Anak Pendek (Stunting) dan Intervensi untuk Mencegah Terjadinya Stunting (Suatu Kajian Kepustakaan). *Jurnal Kesehatan Komunitas*, 2(6), 254–261. <https://doi.org/10.25311/jkk.vol2.iss6.85>.
- Mohammed, M., Khan, M. B., & Bashie, E. B. M. 2016. Machine learning: Algorithms and applications. In *Machine Learning: Algorithms and Applications* (Issue July). <https://doi.org/10.1201/9781315371658>.
- Refaeilzadeh, P., Tang, L., & Liu, H. 2008. Cross-Validation. *Arizona State University*, 25, 1–12. <https://doi.org/10.1159/000331070>.
- Revelle, W. 2020. *psych: Procedures for Psychological, Psychometric, and Personality Research* (2.0.12). website: <https://personality-project.org/r/psych/> <https://personality-project.org/r/psych-manual.pdf> (Akses Juni 06, 2021).

- Riyadi, S., & Sukarmin. 2013. *Asuhan Keperawatan pada Anak*. Graha Ilmu. Yogyakarta: Graha Ilmu.
- Safitri, W. R. 2016. Analisis korelasi pearson dalam menentukan hubungan antara kejadian demam berdarah dengue dengan kepadatan penduduk di kota surabaya pada tahun 2012 - 2014. *Jurnal Ilmu Kesehatan Masyarakat Universitas Airlangga*, 9.
- Sakti, E. S. 2018. Buletin Stunting. Kementerian Kesehatan Republik Indonesia, 301(5), 1163–1178.
- Santra, A. K., & Christy, C. J. 2012. Genetic Algorithm and Confusion Matrix for Document Clustering. *International Journal of Computer Science Issues*, 9(1), 322–328.
- Saputro, B. I. 2017. Penerapan Sistem Klasifikasi Perpustakaan Arkeologi di Perpustakaan Balai Arkeologi Daerah Istimewa Yogyakarta. *Berkala Ilmu Perpustakaan Dan Informasi*, 13(2), 107. <https://doi.org/10.22146/bip.23453>.
- Smola, A., & Vishwanathan, S. V. N. 2009. Introduction To Machine Learning. University Press, Cambridge, I(July), 213. <https://doi.org/10.1007/978-1-62703-748-8-7>.
- Subrata, G. 2009. Klasifikasi bahan pustaka. 1–13. Pustakawan Perpustakaan UM.
- Vickery, R. 2018. *Feature Transformation for Machine Learning, a Beginners Guide*. di medium.com. website: <https://medium.com/vickdata/four-feature-types-and-how-to-transform-them-for-machine-learning-8693e1c24e80> (Akses Oktober 07, 2020).
- Wickham, H. 2020. *readr: Read Rectangular Text Data*. 1.4.0. website: <https://readr.tidyverse.org>, <https://github.com/tidyverse/readr> (Akses Juni 06, 2021).
- Yadav, S., & Shukla, S. 2016. Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. Proceedings - 6th International Advanced Computing Conference, IACC 2016, Cv, 78–83. <https://doi.org/10.1109/IACC.2016.25>.