

**LAPORAN HASIL PENELITIAN
TAHUN ANGGARAN 2020**

JUDUL PENELITIAN
**Model Prediksi *Support Vector Machine* pada Penggunaan Kontrasepsi Hormonal dan
Non Hormonal**

Oleh:
Ria Dhea Layla N.K (2009079002/19900709 20180201 2 228)



KEMENTERIAN AGAMA
LEMBAGA PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT (LP2M)
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2020

DAFTAR ISI

DAFTAR ISI	2
ABSTRAK	3
BAB I PENDAHULUAN	
1.1 Latar Belakang	4
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	7
BAB II TINJUAN PUSTAKA	
2.1 <i>Support Vector Machine</i> (SVM)	8
2.2 Akurasi, Sensitifitas dan Spesifisitas.....	11
2.3 Alat Kontrasepsi pada Program KB dan Faktor yang Mempengaruhi.....	12
BAB III METODOLOGI PENELITIAN	
3.1 Metode dan Teknik Pengumpulan Data	15
3.2 Deskripsi Variabel.....	15
3.3 Rencana Penelitian	16
BAB IV HASIL PENELITIAN	
4.1 Analisis <i>Support Vector Machine</i> (SVM).....	17
4.2 Hasil Akurasi, Sensitifitas dan Spesifitas	19
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	22
5.2 Saran.....	22
DAFTAR PUSTAKA	23
LAMPIRAN	25

ABSTRAK

Di Indonesia setiap 3 menit, 1 anak balita di Indonesia meninggal dunia dan setiap jam, 1 wanita meninggal dunia ketika melahirkan atau karena masalah kehamilan. Hal ini menunjukkan bahwa peningkatan kesehatan bagi ibu berjalan lambat dalam beberapa tahun terakhir. Program yang dibuat oleh pemerintah Indonesia salah satunya yaitu program KB atau yang dikenal dengan Keluarga Berencana. Usaha program KB yang paling utama yaitu pencegahan resiko kesehatan terkait terutama kesehatan pada wanita dan mengurangi resiko kematian pada bayi (World Health Organization, 2018). Jumlah kematian ibu (AKI) melahirkan di Indonesia masih terbilang tinggi yaitu 132-323 per 100.000 kelahiran ibu (PKBI, 2017). Program KB yang dilakukan oleh pemerintah yaitu penggunaan alat kontrasepsi sebagai bentuk fasilitas yang diberikan. *Big data* dapat dianalisis dengan *machine learning*, dimana terdapat beberapa macam metode *machine learning* yang dapat digunakan dalam perkembangan ilmu pengetahuan. Metode SVM merupakan salah satu pengembangan *machine learning* yang dapat memprediksi dengan cara mengklasifikasikan atau mengelompokkan sebuah data berskala besar (*big data*) sehingga dapat memberikan kesimpulan yang tepat. Pada penelitian ini metode SVM digunakan untuk memprediksi model kontrasepsi yang difasilitasi oleh pemerintah. Model prediksi alat kontrasepsi hormonal dan non-hormonal dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM) yaitu 33,87%. Nilai ketepatan klasifikasi dengan menggunakan metode *Support Vector Machine* (SVM) yaitu 41,677% dengan kata lain kesalahan performa model yang terbentuk dengan menggunakan metode SVM jauh dari nilai akurasi sebesar 100%.

Kata Kunci: Kesehatan Ibu, Keluarga Berencana, Kontrasepsi Hormonal, Kontrasepsi Non Hormonal, *Support Vector Machine*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di Indonesia peningkatan kesehatan bagi ibu berjalan lambat dalam beberapa tahun terakhir. Padahal peningkatan kesehatan tersebut merupakan Tujuan Pembangunan Milenium (MDG) kelima yang diprogram oleh pemerintah. Setiap 3 menit, 1 anak balita di Indonesia meninggal dunia dan setiap jam, 1 wanita meninggal dunia ketika melahirkan atau karena masalah kehamilan (UNICEF, 2015). Salah satu program yang dibuat oleh pemerintah Indonesia dalam mengatasi peningkatan kesehatan bagi ibu dan anak yaitu program Keluarga Berencana atau yang dikenal dengan KB.

Program KB mulai diselenggarakan pada akhir tahun 1970-an. Program KB merupakan program merencanakan program kehamilan untuk membentuk keluarga sehat dan sejahtera yaitu dengan membatasi jumlah kelahiran. Program ini diawasi langsung oleh pemerintah melalui Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN). Kesehatan wanita di Indonesia masih perlu diperhatikan. Apalagi jumlah kematian ibu (AKI) melahirkan di Indonesia masih terbilang tinggi yaitu 132-323 per 100.000 kelahiran ibu (PKBI, 2017). Usaha program KB yang paling utama yaitu pencegahan resiko kesehatan terkait terutama kesehatan pada wanita dan mengurangi resiko kematian pada bayi (World Health Organization, 2018). Selain itu manfaat lain program KB membantu mencegah HIV/AIDS, memberdayakan masyarakat dan meningkatkan pendidikan, mengurangi kehamilan remaja, dan memperlambat pertumbuhan populasi. Program KB dilakukan dengan penggunaan alat kontrasepsi seperti kondom, IUD, Spiral, dan sebagainya.

Selain itu, Indonesia merupakan negara kepulauan yang terdiri dari 17.504 pulau dan terdiri dari 1.340 suku bangsa menurut sensus BPS tahun 2010. Indonesia menempati peringkat keempat sebagai negara dengan jumlah populasi terpadat di dunia menurut badan sensus dunia yaitu 264.935.824 penduduk. Proyeksi pertumbuhan penduduk di Indonesia menurut (Badan Pusat Statistik, 2018) akan mengalami peningkatan pada skenario A, yaitu sebesar 0,94%. Hal ini menjadi perhatian pemerintah dalam mengatasi jumlah penduduk di Indonesia.

Alat kontrasepsi yang direkomendasikan oleh BKKBN dibedakan menjadi dua macam yaitu alat kontrasepsi hormonal (pil, suntik, susuk, dan implant) dan alat kontrasepsi fisik (spiral IUD, kondom, diafragma, MOW atau metode operatif wanita). Alat kontrasepsi fisik seperti IUD paling sedikit menimbulkan keluhan dibanding alat kontrasepsi hormonal seperti

suntik dan pil. Namun, masyarakat masih belum banyak mengetahui alat kontrasepsi seperti IUD, diafragma maupun MOW (Buletin Kesehatan Reproduksi, 2013).

Support Vector Machine atau yang dikenal dengan SVM merupakan salah satu metode *machine learning* yang digunakan. SVM pertama kali diperkenalkan oleh Hava Siegelmann dan Vladimir Vapnik. Metode ini dapat digunakan untuk data dengan skala tinggi atau *high dimensional data* sehingga dapat menghasilkan keputusan yang tepat dengan penerapan *Gaussian kernel*. SVM merupakan model *supervised learning* yang berupa klasifikasi atau analisis regresi (Gholami & Fakhari, 2017). Apabila variabel responnya berupa kategori model yang terbentuk adalah *Support Vector Classification*. Namun apabila variabel responnya berupa rasio maka akan menghasilkan model berupa *Support Vector Regression*. Metode SVM memiliki dibedakan menjadi dua kategori, *Support Vector Classification* dan *Support Vector Regression*.

Sebuah penelitian di Nepal, pengaruh kelahiran di negara tersebut dipengaruhi oleh alasan sosial yaitu memiliki banyak anak merupakan simbol kemakmuran sosial dan ekonomi (Adhikari, 2010). (Oddens, 1999) dalam penelitiannya mengenai kepuasan wanita terhadap penggunaan alat kontrasepsi sebagai bentuk upaya pemerintah dalam mengatasi ledakan penduduk di Jerman Barat menggunakan variabel usia, daerah, status perkawinan, pendidikan, status pekerjaan, dan jumlah anak, jumlah kematian pada anak. Metode SVM pada penelitian ini diharapkan mampu memprediksi alat kontrasepsi hormonal dan non hormonal dengan mengklasifikasikan pengguna alat kontrasepsi tersebut faktor tingkat pendidikan, status pekerjaan, status perkawinan, usia istri, usia produktif, anak hidup, dan kelompok usia balita. Sehingga nantinya dapat digunakan sebagai bahan evaluasi pemerintah dan BKKBN dalam upaya peningkatan kesehatan bagi ibu dan bayi serta menciptakan keluarga sehat dan sejahtera di Indonesia.

(Asria, Mousannif, Moatassim, & Noel, 2016) menggunakan beberapa algoritma *machine learning* untuk memprediksi resiko dan diagnosis kanker payudara pada wanita. Algoritma *machine learning* yang digunakan yaitu C4.5, SVM, Naïve Bayes, dan k-NN dengan membandingkan tingkat akurasi dan tingkat kesalahan yang dilakukan oleh metode tersebut. Hasilnya SVM memberikan hasil yang lebih baik dibanding tiga metode yang digunakan. SVM memiliki tingkat akurasi yang tinggi yaitu 97,3% dengan tingkat kesalahan yaitu 0,02% untuk memprediksi resiko dan diagnosis kanker payudara. Pada penelitian tersebut menggunakan klasifikasi jenis kanker payudara jinak dan ganas dengan 11 variabel prediktor dan sampel sebanyak 699 penderita kanker payudara.

(Yu, Liu, Valdez, Gwinn, & Khoury, 2010) mengaplikasikan model SVM untuk memprediksi penyakit umum pada kasus diabetes dan pre-diabetes. Sampel yang digunakan merupakan sampel kecil namun memiliki banyak variabel atau berdimensi tinggi. Pada penelitian tersebut menggunakan data yang berasal dari badan kesehatan nasional tahun 1999-2004 di U.S. Variabel yang digunakan merupakan riwayat keluarga, usia, jenis kelamin, ras dan etnik, berat badan, tinggi badan, ukuran lingkaran perut, BMI riwayat hipertensi, aktivitas fisik, kebiasaan merokok, kebiasaan mengonsumsi alkohol, pendidikan dan pendapatan keluarga. Area di bawah kurva atau yang dikenal dengan AUC (*Area Under The Curve*) yang merupakan karakteristik ROC atau kurva penerimaan dan beberapa evaluasi secara statistik (*sensitivity, specificity*, nilai prediksi positif dan negatif) digunakan untuk mengevaluasi metode SVM. Hasilnya teknik SVM memberikan solusi yang lebih efisien dibanding metode statistik klasik seperti regresi logistik.

Peranan pemerintah dalam mendukung program Keluarga Berencana sangat diperlukan karena menyangkut kesejahteraan kesehatan pada wanita terutama angka kematian ibu melahirkan (AKI). Hal tersebut merupakan salah satu indikator dalam menggambarkan kesejahteraan masyarakat di suatu daerah. Metode SVM, sebagai salah satu metode *machine learning*, akan mengelompokkan serta memprediksi penggunaan alat kontrasepsi hormonal dan non hormonal dengan faktor yang mempengaruhi penggunaan alat kontrasepsi yang disarankan. Sehingga pemerintah dapat mengevaluasi program KB berdasarkan alat kontrasepsi serta memberikan sosialisasi dan edukasi terutama bagi kesehatan wanita agar tepat sasaran.

1.2 Rumusan Masalah

Rumusan masalah yang dibahas pada penelitian ini berdasarkan latar belakang yang telah dijelaskan sebagai berikut

1. Bagaimana model prediksi alat kontrasepsi hormonal dan non hormonal dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM)?
2. Bagaimana nilai ketepatan klasifikasi dengan menggunakan metode *Support Vector Machine* (SVM)?

1.3 Tujuan

Tujuan yang dicapai berdasarkan rumusan masalah sebagai berikut

1. Mengetahui prediksi alat kontrasepsi hormonal dan non hormonal dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM)

2. Mengetahui nilai ketepatan klasifikasi dengan menggunakan metode *Support Vector Machine* (SVM)

1.4 Manfaat Penelitian

Manfaat penelitian sesuai dengan tujuan penelitian dibedakan berdasarkan kepentingan beberapa pihak, yaitu

1. Bagi Penulis
 - a. Menambah wawasan mengenai pentingnya kesehatan reproduksi wanita untuk mengurangi resiko Angka Kematian Ibu dan Bayi
 - b. Menambah wawasan dan pengetahuan dibidang pengelompokkan, terutama *machine learning* sehingga dapat mengembangkan metode yang digunakan serta *tools* yang digunakan
2. Bagi Pemerhati Matematika

Sebagai tambahan pengetahuan dan wawasan dalam keilmuan matematika terutama dalam perkembangan *big data* dan cara menganalisisnya sehingga memberikan kontribusi yang nyata.
3. Bagi Lembaga
 - a. Memberikan wawasan sebagai bahan edukasi terutama bagi kesehatan wanita agar mengurangi resiko Angka Kematian Ibu dan Bayi terutama bagi pasangan suami istri maupun remaja
 - b. Sebagai bahan evaluasi kepada pemerintah pentingnya masyarakat mendukung program Keluarga Berencana untuk kesehatan wanita serta membentuk keluarga sehat dan sejahtera
 - c. Sebagai tambahan bahan literatur dan sebagai sarana pengembangan wawasan keilmuan khususnya di bidang analisis pengelompokkan atau klasifikasi

BAB II

TINJAUAN PUSTAKA

2.1 *Support Vector Machine* (SVM)

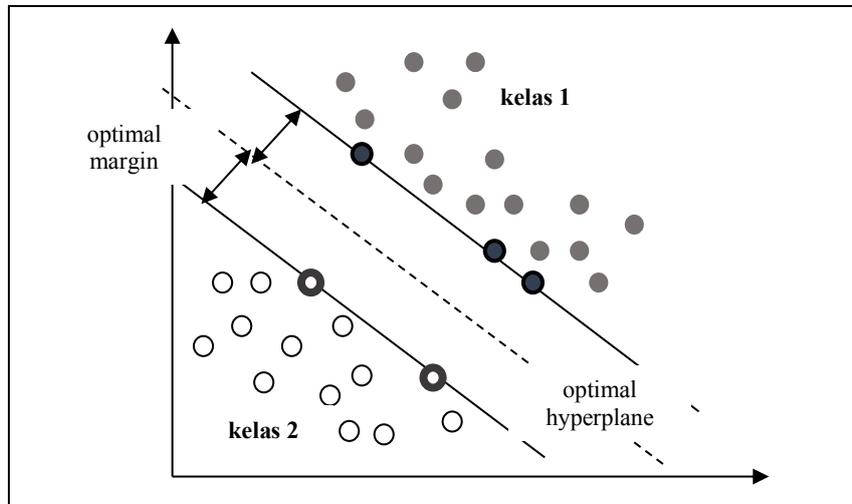
Support Vector Machine (SVM) merupakan salah satu metode *machine learning* baru untuk melakukan prediksi yang dibuat awal tahun 1990-an sebagai penyelesaian *non linear* baik untuk klasifikasi maupun regresi. Metode SVM ini telah mendapat perhatian dari para ilmuwan maupun praktisi untuk menyelesaikan masalah nyata dalam bidang kesehatan, *finance*, cuaca, mesin, minyak bumi, dan lain-lain. Ide dasar dari SVM yaitu membentuk sebuah $n - 1$ dimensi yang dibatasi dengan *hyperplane* untuk membedakan dua kelas pada sebuah ruang berdimensi n (Yu, Liu, Valdez, Gwinn, & Khoury, 2010). SVM dapat memiliki fungsi pemisah atau *classifier*. Fungsi pemisah tersebut memisahkan dua set dari dua kelas yang berbeda dapat dibedakan menjadi kelas 1 dan -1. SVM menggunakan fungsi kernel (seperti fungsi *linear*, *polynomial*, *sigmoid* dan basis fungsi radial atau *Radial Basis Function* (RBF)) yang berfungsi untuk menambahkan lebih banyak dimensi ke ruang dimensi rendah sehingga pemisah atau *classifier* tersebut dapat dipisahkan ke dalam ruang berdimensi tinggi.

Data yang digunakan dalam observasi menggunakan SVM merupakan sebuah vektor berdimensi n . Misalnya terdapat dua variabel pada sebuah set data maka akan terbentuk dua ruang dimensi, pembatas *hyperplane* akan membentuk garis lurus (dimensi 1) yang membagi ruang menjadi dua bagian. Namun, ketika lebih dari dua dimensi yang digunakan, maka SVM mencari pemisah *hyperplane* optimum yang disebut pembatas *hyperplane* maksimal, dengan kata lain *hyperplane* optimal atau terbaik berfungsi sebagai pemisah dua kelas pada ruang input. Jarak antara *hyperplane* dan titik data terdekat pada tiap sisi (disebut juga *support vectors*) yaitu jarak maksimum. Pada kenyataannya terkadang terdapat titik data yang terletak dalam garis dan tidak mudah untuk dipisahkan, sehingga SVM dapat mengatasi masalah tersebut.

Hyperplane didefinisikan sebagai fungsi keputusan yang menghasilkan margin maksimum antara vektor dua kelas (Gambar 2.1). *Hyperplane* diambil dari kumpulan data *training*, x_i ($i = 1, 2, 3, \dots, n$), atau data yang digunakan untuk membentuk model pada *machine learning*. Nilai *hyperplane* dikatakan optimal pada suatu ruang (Cortes & Vapnik, 1995),

$$w \cdot x_i + b_0 = 0 \tag{2.1}$$

dimana, w merupakan vektor berdimensi n dan b merupakan kondisi bias. *Hyperplane* harus memiliki dua kondisi yang spesifik yaitu meminimumkan kesalahan dalam memisahkan kelas dan jarak dari data terdekat dari masing-masing kelas harus maksimal (Gholami & Fakhari, 2017).



Gambar 2.1 Support Vector pada Dua Kelas

Pada klasifikasi data dibagi menjadi dua kelas (Gambar 2.1) yang dapat dinotasikan dengan sisi *hyperplane* sebelah kanan yaitu $y=1$ dan sisi sebelah kiri $y=-1$. Kedua margin tersebut didefinisikan untuk mengontrol data (yang digunakan untuk semua elemen data training),

$$\begin{aligned} w \cdot x_i + b &\geq 1 \text{ jika } y_i = 1 \\ w \cdot x_i + b &\leq -1 \text{ jika } y_i = -1 \end{aligned} \quad (2.2)$$

2.1.1 Algoritma *Hyperplane* Optimal

Kumpulan dari bentuk training data dapat berupa

$$(y_1, x_1), \dots, (y_\ell, x_\ell) \text{ dimana } y_i \in \{-1, 1\} \quad (2.3)$$

Pertidaksamaan (2.2) dan persamaan (2.3) dapat dinyatakan dengan pertidaksamaan dalam bentuk

$$y_i(w \cdot x_i + b) \geq 1 \text{ dimana } i = 1, \dots, \ell \quad (2.4)$$

Hyperplane optimal dapat dinyatakan dengan (Cortes & Vapnik, 1995),

$$w_0 \cdot x_i + b_0 = 0 \quad (2.5)$$

pada persamaan (2.5) akan memisahkan data training dengan margin maksimal untuk menentukan arah $w/|w|$ dimana jarak antara proyeksi atau letak training vektor dari dua kelas adalah maksimal (lihat Gambar 2.1). Jarak $d(w, b)$ yaitu

$$d(w, b) = \min_{\{x:y=1\}} \frac{x \cdot w}{|w|} - \max_{\{x:y=-1\}} \frac{x \cdot w}{|w|} \quad (2.6)$$

Sehingga, *hyperplane* optimal (w_0, b_0) merupakan pernyataan untuk memaksimumkan jarak (2.6). Berdasarkan persamaan (2.6) dan (2.4), maka

$$d(w_0, b_0) = \frac{2}{|w_0|} = \frac{2}{\sqrt{w_0 \cdot w_0}} \quad (2.7)$$

Artinya, persamaan (2.7) merupakan *hyperplane* optimal yang mana nilainya dapat meminumkan $w \cdot w$ dengan syarat persamaan (2.4). Vektor x_i untuk $y_i(w \cdot x_i + b)$ disebut *support vectors*.

Persamaan *hyperplane* optimal (2.5) yang membagi kumpulan data training (2.3) memiliki fungsi untuk meminumkan kendala

$$\pi = w \cdot w, \quad (2.8)$$

dengan subyek fungsi kendala pada persamaan (2.4).

Teknik optimisasi pada persamaan tersebut dapat diselesaikan dengan *Langrage Multipliers*

$$L(w, b, \alpha) = \frac{1}{2} w \cdot w - \sum_{i=1}^{\ell} \alpha_i [y_i(w \cdot x_i + b) - 1] \quad (2.9)$$

dimana, $\alpha^T = (\alpha_1, \dots, \alpha_{\ell})$ merupakan vektor non-negatif *Langrage Multipliers* (bernilai nol atau positif) dengan fungsi kendala pada persamaan (2.3) atau *Langrage Multipliers* yang berkorepondensi dengan x_i .

Teknik tersebut dapat disesaikan, pertama dengan meminimkan nilai L terhadap w dan kedua memaksimumkan L terhadap α^T . Berikut modifikasi persamaan $L(w, b, \alpha)$,

$$\left. \frac{\partial L(w, b, \alpha)}{\partial w} \right|_{w=w_0} = (w_0 - \sum_{i=1}^{\ell} \alpha_i y_i x_i) = 0 \quad (2.10)$$

$$\left. \frac{\partial L(w, b, \alpha)}{\partial b} \right|_{b=b_0} = \sum \alpha_i \alpha_i y_i = 0 . \quad (2.11)$$

Sehingga, dari persamaan (2.10) diperoleh

$$w_0 = \sum_{i=1}^{\ell} \alpha_i y_i x_i \quad (2.12)$$

dengan mensubstitusi persamaan (2.12) dab (2.10) didapatkan

$$\begin{aligned} W(\alpha) &= \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} w_0 \cdot w_0 \\ &= \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \end{aligned} \quad (2.13)$$

Maka, akan menghasilkan

$$\alpha_i \geq 0, \sum_{i=1}^{\ell} \alpha_i y_i = 0 \quad (2.14)$$

Solusi dari masalah tersebut akan menghasilkan beberapa α_i dengan banyak nilai nol. Sedangkan, data yang berkorespondensi dengan α_i yang tidak nol merupakan *support vector*, yaitu jarak terdekat dengan *hyperplane*.

2.1.2 Soft Margin

Proses optimisasi tidak diselesaikan apabila w dan b tidak memenuhi persamaan (2.4). Penjabaran yang telah dilakukan merupakan penjabaran berdasarkan asumsi jika kedua kelas dapat dipisah sempurna oleh *hyperplane*. Tetapi, pada umumnya kedua kelas tidak dapat terpisah secara sempurna sehingga proses optimisasi tidak dapat dilakukan.

Pada persamaan (2.4) dapat dimodifikasi dengan memasukkan beberapa variabel non-negatif tambahan atau *slack variable* $\xi_i \geq 0$ dimana $i = 1, \dots, \ell$ (Cortes & Vapnik, 1995)

$$\Phi(\xi) = \sum_{i=1}^{\ell} \xi_i^{\sigma} \quad (2.15)$$

jika $\sigma > 0$ bernilai kecil, maka fungsi kendala

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \text{ dimana } i = 1, \dots, \ell \quad (2.16)$$

$$\xi_i \geq 0 \text{ dimana } i = 1, \dots, \ell. \quad (2.17)$$

Nilai σ yang cukup kecil pada persamaan (2.15) merupakan banyaknya kesalahan atau *error* pada data training. Sehingga, untuk meminimumkan nilai *error*

$$\frac{1}{2} w^2 + C \left(\sum_{i=1}^{\ell} \xi_i^{\sigma} \right) \quad (2.18)$$

dimana, parameter C berfungsi untuk mengontrol *tradeoff* antar margin dan kesalahan klasifikasi. Apabila nilai C semakin besar maka semakin besar pula kesalahan untuk tiap kelas. Jika data training dapat dipisahkan tanpa *error* maka *hyperplane* yang dibuat dapat tepat sesuai dengan *hyperplane* optimal.

2.2 Akurasi, Sensitifitas dan Spesifisitas

Salah satu metode yang digunakan untuk mengukur ketepatan ukuran klasifikasi yaitu *Apparent Error Rate* (APER). APER merupakan metode yang dapat digunakan untuk prosedur klasifikasi apapun dan tidak bergantung pada populasi induknya. Semakin tinggi nilai akurasi klasifikasi artinya semakin baik performa pada teknik klasifikasi (Johnson & Wichern, 2008).

Tabel 2.1 *Apparent Error Rate* (APER)

		Kelas Prediksi	
		Positif (1)	Negatif (2)
Kelas Aktual	Positif (1)	TN <i>(True Negative)</i>	FN <i>(False Negative)</i>
	Negatif (2)	FP <i>(False Positive)</i>	TP <i>(True Positive)</i>

dimana,

TN : banyak observasi dari kelas 1 yang tepat diklasifikasikan sebagai kelas 1
 FP : banyak observasi dari kelas -1 yang tidak tepat diklasifikasikan sebagai kelas 1
 FN : banyak observasi dari kelas 1 yang tidak tepat diklasifikasikan sebagai kelas -1
 TP : banyak observasi dari kelas -1 yang tepat diklasifikasikan sebagai kelas -1
 dengan nilai APER yaitu,

$$APER (\%) = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (2.19)$$

Nilai akurasi pada data yang tidak seimbang seringkali mengalami ketidaksesuaian sebagai ukuran kinerja. Oleh sebab itu, kurva ROC (*Receiver Operating Characteristic*) lebih dipertimbangkan (Agresti, 2007) dibandingkan APER dalam mengukur akurasi dalam klasifikasi. Kurva ROC memiliki nilai sensitivitas dan spesifisitas. Sensitivitas merupakan akurasi pada kelas positif sedangkan spesifisitas merupakan akurasi pada kelas negatif. Nilai akurasi sensitivitas bergantung pada jumlah *true positive* (TP) dan *true negative* (TN). Sedangkan nilai spesifisitas bergantung pada jumlah dari *true negative* (TN) dan *false positive* (FP).

$$\text{Sensitivitas} = \frac{TP}{TP+FN} \quad (2.20)$$

$$\text{Spesifisitas} = \frac{TN}{TN+FP} \quad (2.21)$$

2.3 Alat Kontrasepsi pada Program KB dan Faktor yang Mempengaruhi

Pemerintah Indonesia untuk mencapai sasaran *Millenium Development Goals* (MDGs) yaitu Angka Kematian Ibu (AKI) sebesar 102 per 100.000 kelahiran hidup dan Angka Kematian Bayi (AKB) menjadi 23 per 1000 kelahiran hidup pada tahun 2015 memerlukan usaha percepatan yang besar dan kerja keras pada saat ini karena masih meningkatnya AKI dan AKB di Indonesia (Kementerian Kesehatan Republik Indonesia, 2010). Pemerintah juga telah menganggarkan bantuan operasional kesehatan di Puskesmas baik Kabupaten maupun Kota sebagai bentuk perhatian. Salah satu upaya pencegahan dalam penurunan AKI dan AKB merupakan program perencanaan dan pencegahan (Kementerian Kesehatan Republik Indonesia, 2010). Pasangan suami dan istri seharusnya meningkatkan persiapan komplikasi pada saat kehamilan termasuk merencanakan pemakaian alat maupun obat kontrasepsi pasca melahirkan. Perencanaan akan berjalan baik apabila suami dan istri memiliki pengetahuan mengenai tanda bahaya kehamilan, jadwal imunisasi, persalinan dan nifas dan lain-lain.

Program Keluarga Berencana atau yang dikenal dengan KB merupakan program yang dibuat pemerintah pada tahun 1970-an. Tujuan awal dari program ini yaitu mewujudkan masyarakat Indonesia yang sehat dan sejahtera. Pada buletin Jendela Data dan Informasi yang diterbitkan oleh (Dinas Kesehatan, 2013), program KB telah diatur dalam Undang-Undang

Nomor 29 Tahun 2009 mengenai perkembangan kependudukan dan pembangunan keluarga yang menyatakan bahwa keluarga berencana merupakan upaya untuk mengatur kelahiran anak, jarak, dan usia ideal melahirkan, mengatur kehamilan, melalui promosi, perlindungan, dan bantuan sesuai hak reproduksi untuk menunjukkan keluarga berkualitas. Masalah kependudukan dan pembangunan keluarga juga telah dibahas dalam seminar internasional puplasi dan pengembangan (ICPD) di Kairo, Mesir yaitu memiliki orientasi mengutamakan kesehatan reproduksi perorangan dengan menghormati hak reproduksi setiap individu atau manusia sebagai subyek.

Jenis alat kontrasepsi di Indonesia dibagi menjadi dua, yaitu hormonal dan non hormonal. Kontrasepsi hormonal merupakan metode dalam mengontrol kehamilan dengan system endrokrin, misalnya penggunaan pil, suntik, implant maupun alat kontrasepsi dalam Rahim (AKDR). Kontrasepsi non hormonal merupakan alat kontrasepsi yang tidak mengandung hormon sehingga penggunaan alat kontrasepsi ini tidak mempengaruhi kondisi hormonal dalam tubuh misalnya kondom, IUD, operasi medis (steril pada wanita atau pria).

Berikut faktor penggunaan alat kontrasepsi baik alat kontrasepsi hormonal maupun non hormonal (Buletin Kesehatan Reproduksi, 2013)

1. Tingkat Pendidikan

Menurut Pusat Data dan Informasi dari kemenrian kesehatan menyatakan bahwa masyarakat yang memiliki tingkat pendidikan tinggi lebih menggunakan alat kontrasepsi dan lebih mengetahui mengenai reproduksi pasangannya. Tingkat pendidikan dalam kasus ini dibagi menjadi SD, SMP, SMA dan lebih dari SMA

2. Status Pekerjaan

Status pekerjaan digolongkan menjadi kepala keluarga yang bekerja dan tidak bekerja

3. Status Perkawinan

Status perkawinan digolongkan menjadi golongan kawin dan duda/janda/belum kawin

4. KK Menurut Anak Hidup

Kepala Keluarga menurut anak hidup yaitu banyaknya anak yang dimiliki oleh kepala keluarga dan masih hidup yaitu kurang dari 2 tahun dan lebih dari sama dengan 2 tahun

5. Alasan ber-KB

Alasan ber-KB merupakan alasan pasangan tidak melakukan program KB. Alasan tersebut diantaranya sedang hamil, alasan fertilitas, tidak menyetujui KB, tidak tahu mengenai program KB, takut efek samping, pelayanan KB jauh, dan tidak mampu/mahal

6. KK Kelompok Usia Produktif

Kepala Keluarga kelompok usia produktif yaitu pasangan yang menikah di usia produktif yang dibagi menjadi usia 16 sampai 21 tahun dan 22 sampai 59 tahun

7. KK Kelompok Lansia

Kepala Keluarga kelompok lansia merupakan pasangan atau kepala keluarga yang sudah memiliki usia diatas 59 tahun.

BAB III

METODOLOGI PENELITIAN

3.1 Metode dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini menggunakan data dari tiga provinsi berdasarkan program yang dikembangkan oleh pemerintah dalam perencanaan keluarga atau keluarga berencana yang dikenal dengan *Development Program Implementation Study* (DPIS). (Warwick, 1986) menggunakan empat provinsi tentang kasus peranan pemerintah dalam program perencanaan keluarga dan pilihan masyarakat mengenai program tersebut. Sehingga fokus penelitian ini yaitu, Jawa Timur (program perencanaan keluarga dianggap berhasil dilakukan oleh pemerintah), Sumatra Selatan (kurang berhasil dilakukan oleh pemerintah), dan Jawa Barat (terdapat kasus campuran yaitu berhasil dan tidak berhasil program perencanaan keluarga oleh pemerintah). Data yang akan digunakan merupakan data skunder yang diambil dari Badan Keluarga Berencana Nasional.

3.2 Deskripsi Variabel

Alat kontrasepsi di Indonesia dibedakan menjadi dua jenis, yaitu alat kontrasepsi hormonal dan alat kontrasepsi non hormonal. Pada penelitian ini variabel respon yang digunakan dibedakan menjadi dua kategori yaitu jenis alat kontrasepsi hormonal dan non hormonal. Sehingga, skala data variabel respon yaitu nominal (Hormonal diberi koding 1 dan Non Hormonal diberi koding 2). Sedangkan banyaknya variabel prediktor 21 faktor, yaitu

Tabel 3.1 Deskripsi Variabel

Variabel Prediktor	Nama Variabel	Skala	Keterangan
X1	1: Jawa Timur 2: Sumatera Selatan 3: Jawa Barat	Nominal	Provinsi
X2	Tidak Tamat SD	Rasio	Tingkat Pendidikan
X3	SD-SMP		
X4	Tamat SLTA		
X5	Tamat PT/Akademi		
X6	Bekerja	Rasio	Status Pekerjaan
X7	Tidak Bekerja		
X8	Kawin	Rasio	Status Perkawinan
X9	Duda/Janda/Belum Kawin		

Lanjutan Tabel 3.1 Deskripsi Variabel

Variabel Prediktor	Nama Variabel	Skala	Keterangan
X10	<2 tahun	Rasio	KK Menurut Anak Hidup
X11	>= 2 tahun		
X12	16-21 tahun	Rasio	ProduktifKK Kelompok Usia
X13	22-59 tahun		
X15	Sedang Hamil	Rasio	Alasan ber-KB
X16	Fertilitas		
X17	Tidak Menyetujui KB		
X18	Tidak Tahu Tentang KB		
X19	Takut Efek Samping		
X20	Pelayanan KB Jauh		
X21	Tidak Mampu/Mahal		

3.3 Rencana Pembahasan

Adapun rencana pembahasan yang akan dilakukan dalam pelaksanaan penelitian ini sebagai berikut

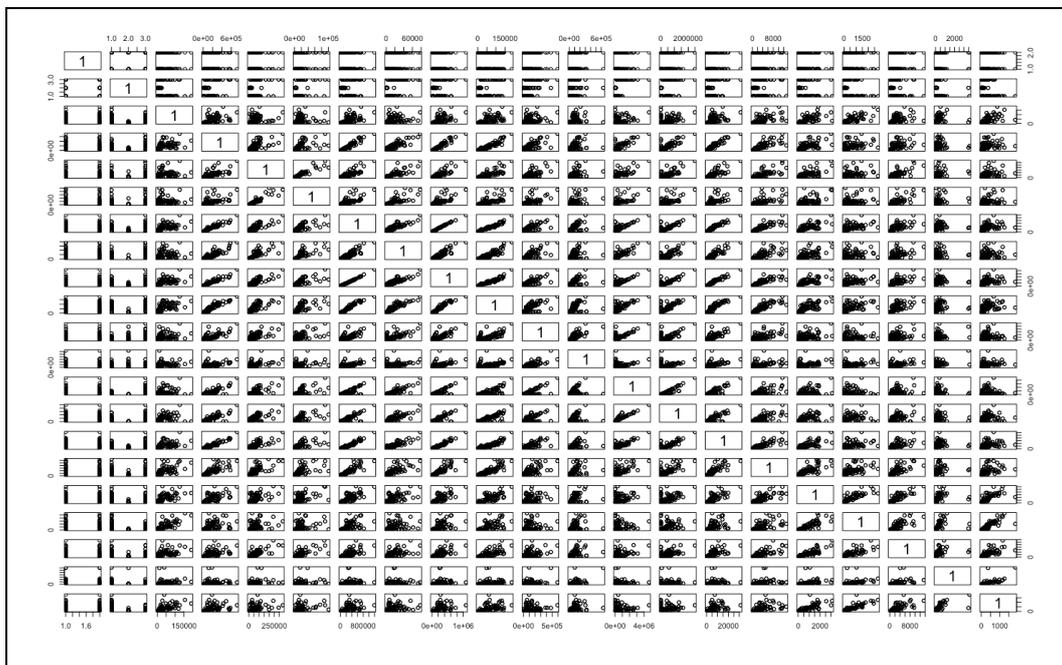
1. Hal yang pertama kali dilakukan sebelum melakukan analisis pada sebuah data yaitu mendapatkan dan mempersiapkan data
2. Data yang telah didapat kemudian dilakukan *screening* agar tidak menimbulkan bias interpretasi
3. Data yang didapat diolah dengan menggunakan metode *Support Vector Machine* (SVM)
 - a. Memberi koding untuk penggunaan alat kontrasepsi hormonal dan non hormonal. Jika menggunakan alat kontrasepsi hormonal dapat diberi koding 1 dan jika menggunakan alat kontrasepsi non hormonal diberi koding -1
 - b. Setelah memberi koding pada variabel prediktor membagi data menjadi dua bagian yaitu *data training* dan *data testing*. Data training digunakan untuk membentuk model prediksi klasifikasi dengan SVM sedangkan data testing digunakan untuk mengecek kesalahan model prediksi
 - c. Menganalisis dengan menggunakan SVM dan bantuan perangkat lunak *cran R*.
4. Membuat kesimpulan mengenai model prediksi alat kontrasepsi hormonal dan non hormonal dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM).
5. Memeriksa nilai ketepatan klasifikasi dari hasil prediksi dengan menggunakan metode *Support Vector Machine* (SVM)

BAB IV

HASIL PENELITIAN

4.1 Analisis *Support Vector Machine* (SVM)

Model prediksi alat kontrasepsi hormonal dan non hormonal pada penelitian ini dianalisis dengan menggunakan salah satu metode *machine learning*, yaitu *Support Vector Machine* (SVM). Jumlah variabel prediktor atau variabel bebas yaitu sebanyak 2, sedangkan variabel terikat atau variabel dependen memiliki dua kelas, yaitu hormonal dan non-hormonal. Beberapa bentuk kelas yang terbentuk sebelum dilakukan pengelompokan dengan SVM ditunjukkan pada Gambar 4.1. Sehingga proses pengelompokan dengan dengan SVM dapat digunakan untuk membantu pengelompokan kelas.



Gambar 4.1 Plot Data Alat Kontrasepsi Hormonal dan Non-Hormonal

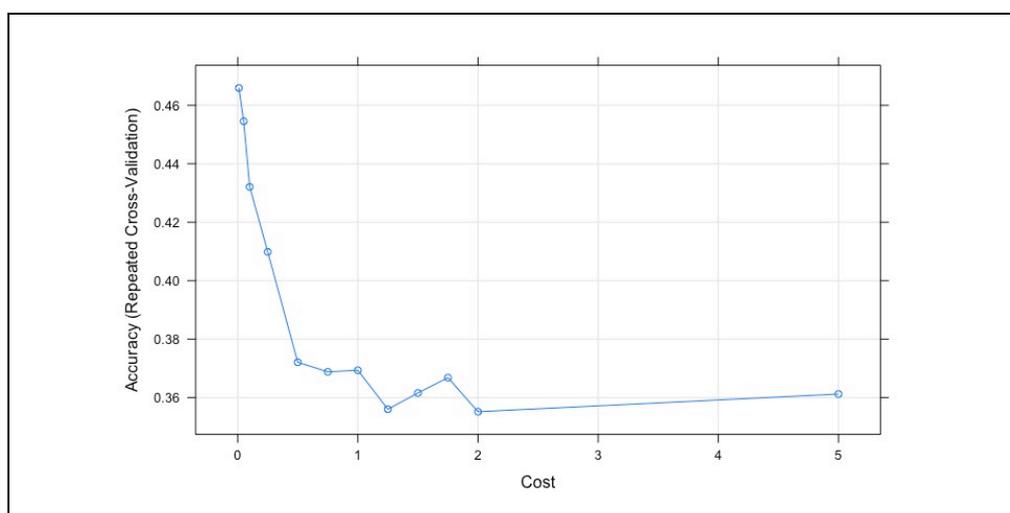
Sebelum dilakukan analisis dengan SVM, terlebih dahulu data dibagi menjadi dua, yaitu data training dan data testing. Data training atau yang disebut juga dengan *learning data* digunakan sebagai pembentukan model pada metode *machine learning* atau dapat digunakan untuk melihat akurasi model algoritma yang sedang digunakan. Sedangkan data *testing* digunakan untuk mengetahui kinerja model yang dihasilkan dengan kenyataan. Prediksi model dilakukan dengan bantuan *packages 'caret'* pada *software R-Gui*. Hasil analisis dengan metode SVM dapat dilihat pada Tabel 4.1

Tabel 4.1 Hasil Analisis dengan SVM

Support Vector Machines with Linear Kernel	
118 samples	
20 predictor	
2 classes: '1', '2'	
Pre-processing: centered (20), scaled (20)	
Resampling: Cross-Validated (10 fold, repeated 3 times)	
Summary of sample sizes: 107, 106, 107, 106, 106, 106, ...	
Resampling results:	
Accuracy	Kappa
0.3387374	-0.3186043

Tabel 4.1 menunjukkan kekuatan model yang dibangun 0,3387374 serta model yang digunakan merupakan model yang linear yang ditunjukkan juga pada nilai C yaitu lebih dari 0. Nilai C merupakan *cost* yang digunakan untuk menentukan kemungkinan kesalahan klasifikasi. Apabila nilai C semakin tinggi maka semakin tinggi SVM salah dalam mengklasifikasikan titik dan sebaliknya. Nilai Kappa menunjukkan bahwa model yang dihasilkan kurang dari yang diharapkan karena memiliki nilai negatif. Namun, menurut (Delgado & Tibau , 2019) tidak dapat dijadikan acuan melihat performa ukuran klasifikasi.

Gambar 4.2 merupakan plot model akurasi dengan nilai C. Gambar 4.2 menunjukkan nilai C akan konstan ketika berharga 5. Sehingga, hasil nilai C yang diprediksi oleh model yaitu dengan 0,01.



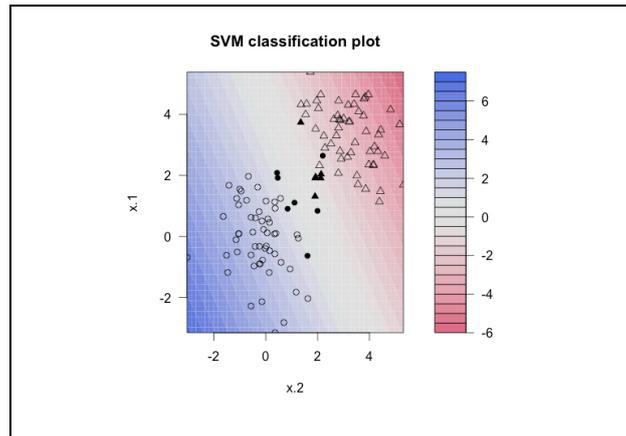
Gambar 4.2 Plot Model Akurasi Data Training

Setelah dilakukan analisis dengan menggunakan SVM, maka variabel yang berpengaruh pada KB hormonal dan non-hormonal dapat dilihat pada tabel 4.2 berikut.

Tabel 4.2 Variabel yang Berpengaruh

	Importance
sdg_hamil	0,7758
Provinsi	0,7586
bekerja	0,7320
tdk_bekerja	0,7291
takut_efek_samping	0,7187
tdk_setuju_kb	0,7074
tdktamatsd	0,7034
fertilitas	0,7022
<=2_anak	0,6892
pelayanan_kb_jauh	0,6813
tamatslta	0,6806
22-59thn	0,6704
>= 2_anak	0,6695
tamatsd.smp	0,6636
tdk_mampu_mahal	0,6601
duda_janda	0,6511
kawin	0,6470
16-21thn	0,6460
tdk_tahu	0,6433
pt_atau_akademi	0,6417

Tabel 4.2 merupakan variabel yang terpenting dari KB hormonal dan non-hormonal. Semakin mendekati 1, maka variabel bebas yang digunakan semakin berpengaruh. Misalnya, variabel alasan KB sedang hamil memiliki pengaruh terhadap KB jenis hormonal dan non-hormonal sebesar 77,58% dan seterusnya. Hasil plot pengklasifikasian variabel alasan KB sedang hamil digambarkan pada Gambar 4.3 berikut.



Gambar 4.3 Plot Klasifikasi SVM

4.2 Akurasi, Sensitifitas, dan Spesifitas SVM

Data testing digunakan untuk mengetahui hasil performa atau kinerja prediksi telah sesuai dengan kenyataan digunakan. Tabel 4.3 merupakan hasil akurasi, sensitifitas dan spesifitas.

Tabel 4.3 Hasil Akurasi, Sensitifitas, dan Spesifitas SVM

Confusion Matrix and Statistics	
test_pred_grid	1 2
1	15 19
2	9 5
Accuracy : 0.4167	
95% CI : (0.2761, 0.5679)	
No Information Rate : 0.5	
P-Value [Acc > NIR] : 0.90329	
Kappa : -0.1667	
McNemar's Test P-Value : 0.08897	
Sensitivity : 0.6250	
Specificity : 0.2083	
Pos Pred Value : 0.4412	
Neg Pred Value : 0.3571	
Prevalence : 0.5000	
Detection Rate : 0.3125	
Detection Prevalence : 0.7083	
Balanced Accuracy : 0.4167	
'Positive' Class : 1	

Tabel 4.3 menunjukkan kesalahan dalam klasifikasi metode SVM, sebesar 0,4167 yang artinya kesalahan klasifikasi model yang terbentuk dalam observasi yaitu 41,677% dengan kata

lain performa model yang terbentuk dengan menggunakan metode SVM jauh dari nilai akurasi sebesar 100%. Semakin nilai akurasi mendekati 100%, maka kesalahan dalam membentuk model semakin rendah. Sensitifitas yang dihasilkan yaitu 0,6250 yang artinya 62,50% data pengguna KB hormonal yang benar diklasifikasikan sebagai pengguna KB hormonal sebesar 62,50%. Sedangkan, sensitifitas merupakan nilai yang menunjukkan sebanyak 0,2083 atau pengguna KB non-hormonal dapat dengan benar diklasifikasikan sebesar 20,83%.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan hasil penelitian yang dilakukan berdasarkan rumusan masalah sebagai berikut.

1. Model prediksi alat kontrasepsi hormonal dan non-hormonal dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM) yaitu 33,87%. Sehingga, model prediksi dengan menggunakan metode SVM kurang baik.
2. Nilai ketepatan klasifikasi dengan menggunakan metode *Support Vector Machine* (SVM) yaitu 41,677% dengan kata lain kesalahan performa model yang terbentuk dengan menggunakan metode SVM jauh dari nilai akurasi sebesar 100%.

5.2 Saran

Saran untuk penelitian selanjutnya yaitu, membandingkan metode *machine learning* yang lain dengan SVM. Sehingga, dapat dibandingkan nilai keakuratan model maupun ketepatan klasifikasinya dengan metode lain.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2018). *Proyeksi Penduduk Indonesia 2015-2045*. Jakarta, Indonesia: Badan Pusat Statistik.
- PKBI. (2017). Dipetik Juli 30, 2019, dari Perkumpulan Keluarga Berencana Indonesia: <https://pkbi.or.id/kematian-ibu-melahirkan-terus-meningkat/>
- World Health Organization. (2018, February 8). *Family Planning Contraception*. Dipetik July 30, 2019, dari www.who.int/news-room/fact-sheets/detail/family-planning-contraception
- Buletin Kesehatan Reproduksi. (2013, November). Situasi Keluarga Berencana di Indonesia. *Pelayanan KB Pasca Persalinan dalam Upaya Mendukung Percepatan Penurunan Angka Kematian Ibu, II*.
- Gholami, R., & Fakhari, N. (2017). Chapter 27: Support Vector Machine: Principles, Parameters, and Applications. Dalam A. Valutkevich (Penyunt.), *Handbook of Neural Computation* (hal. 215-235). London, United Kingdom: Mara Conner. Diambil kembali dari <https://doi.org/10.1016/B978-0-12-811318-9.00027-2>
- Adhikari, R. (2010). Research article Demographic, socio-economic, and cultural factors. *BMC Pregnancy&Childbirth*. Dipetik Juli 30, 2019, dari <https://biomedcentral.com/1471-2393/10/19/>
- Oddens, B. J. (1999). Women's Satisfaction With Birth Control: A Population Survey of Physical and Psychological Effects of Oral Contraceptives, Intrauterine Devices, Condoms, Natural Family Planning, and Sterilization Among 1466 Women. *Elsevier Science*.
- Asria, H., Mousannif, H., Moatassim, H. A., & Noel, T. (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *The 6th International Symposium on Frontiers in Ambient and Mobile Systems (FAMS 2016)*, (hal. 1064-1069). doi:10.1016/j.procs.2016.04.224
- Yu, W., Liu, T., Valdez, R., Gwinn, M., & Khoury, M. J. (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Medical Informatics&Decision Making*. Dipetik Juli 30, 2019, dari <https://biomedcentral.com/1472-6947/10/16/>
- Cortes, C., & Vapnik, V. N. (1995). Support-Vector Networks. 273-297. doi:10.1007/BF00994018
- Johnson, R. A., & Wichern, D. W. (2008). *Applied Multivariate Statistical Analysis, 6th Edition*. Prentice-Hall.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis Second Edition*. New Jersey: Jon Wiley&Sons, Inc.
- Dinas Kesehatan. (2013). *Situasi dan Analisis Keluarga Berencana*. Jakarta: Kementerian Kesehatan RI.
- Kementerian Kesehatan Republik Indonesia. (2010, Februari 3). <http://www.depkes.go.id>. Dipetik Juli 30, 2019, dari <http://www.depkes.go.id/development/site/jkn/index.php?view=print&cid=793&id=untuk-menurunkan-angka-kematian-ibu-dan-kematian-bayi-perlu-kerja-keras>
- Warwick, D. P. (1986, Septtember). The Indonesian Planning Program: Government Influence and Client Choice. *Population and Development Review*, 12, 453-490. Dipetik Juli 30, 2019, dari <http://www.jstor.org/stable/1973219>
- UNICEF. (2015). Dipetik Juli 30, 2019, dari <https://data.unicef.org/>: <https://data.unicef.org/topic/maternal-health/maternal-mortality/>

UNICEF. (2015). Dipetik Juli 30, 2019, dari <https://data.unicef.org/>:
<https://data.unicef.org/topic/maternal-health/maternal-mortality/>

LAMPIRAN

Lampiran 1. Script yang Dilakukan

```
# install.packages("e1071")
library(e1071)
library(caret)
library(devtools)
library(readxl)
SVM_LITAPDIMAS_2020 <- read_excel("Downloads/SVM LITAPDIMAS 2020.xlsx",
                                col_types = c("numeric", "numeric", "numeric",
                                              "numeric", "numeric", "numeric"))
View(SVM_LITAPDIMAS_2020)
dataSVM = SVM_LITAPDIMAS_2020
#using the anyNA() method, which checks for any null values
anyNA(dataSVM)# which means that there are missing values in our dataset.
#checking the summary of our data by using the summary() function
summary(dataSVM)
set.seed(123456)
## split data into a train and test set
intrain = createDataPartition(dataSVM$StatusKB, p= 0.7, list = FALSE)
training = dataSVM[intrain,]
testing = dataSVM[-intrain,]

#checking the dimensions of our training data frame and testing data frame
dim(training);
dim(testing);

# this should be a categorical variable. To convert these to categorical variables, we need to
factorize them:
training$StatusKB = as.factor(training$StatusKB)
testing$StatusKB = as.factor(testing$StatusKB)
```

```

#the traincontrol() method
trctrl = trainControl(method = "repeatedcv", number = 10, repeats = 3)
set.seed(4567)
model = train(StatusKB ~ ., data = training, method = "svmLinear",
              trControl=trctrl,
              preProcess = c("center", "scale"),
              tuneLength = 10)

model
varImp(model)
plot(model)
# The predict() method returns a list, we are saving it in a test_pred variable
test_pred = predict(model, newdata = testing)
test_pred
confusionMatrix(table(test_pred, testing$StatusKB))
# It needs to be put in train() method with tuneGrid parameter.
set.seed(56789)
grid <- expand.grid(C = c(0,0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2,5))
svm_Linear_Grid <- train(StatusKB ~., data = training, method = "svmLinear",
                        trControl=trctrl,
                        preProcess = c("center", "scale"),
                        tuneGrid = grid,
                        tuneLength = 10)

svm_Linear_Grid
plot(svm_Linear_Grid)

```