

# PENGGUNAAN PROPENSITY SCORE STRATIFICATION-SUPPORT VECTOR MACHINE UNTUK MENGESTIMASI EFEK PERLAKUKAN AKTIVITAS OLAHRAGA PADA PENDERITA DIABETES MELITUS\*

Ernawati<sup>1‡</sup>, Bambang Widjanarko Otok<sup>2</sup>, and Sutikno<sup>3</sup>

<sup>1</sup>Department of Sufism and Psychotherapy, UIN Walisongo Semarang, Indonesia, ernawati.offstat@gmail.com

<sup>2</sup>Department of Statistics, ITS Surabaya, Indonesia, bambang\_bw@statistika.its.ac.id

<sup>3</sup>Department of Statistics, ITS Surabaya, Indonesia, sutikno@statistika.its.ac.id

‡corresponding author

**Indonesian Journal of Statistics and Its Applications (eISSN:2599-0802)  
Vol 4 No 3 (2020), 510 - 527**

Copyright © 2020 Ernawati, Bambang Widjanarko Otok, and Sutikno. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## **Abstract**

Randomized Controlled Trial (RCT) is not possible to do in observational studies, mainly in health cases, because it is directly related to human life. Actually, good randomization is needed to make the treatment and control groups have no large differences in the observed variables, so it results from unbiased treatment. One alternative method that is increasingly used in statistical analysis in the field of health is the Propensity Score (PS). If the propensity score had estimated using the SVM method and divided into groups of strata that have a similar propensity score, it is known as the Propensity Score Stratification (PSS-SVM). The purpose of the PSS-SVM is to balance the observed variables between the treatment group and the control group by dividing them into several strata groups so that a balanced trend is obtained or the propensity score is called balance. This eliminates the influence of the confounding variables and unbalance of the treatment and control groups and obtain an unbiased estimation of the treatment effect. In this Research, the PSS-Method applied in case of disease complication in patients with Diabetes Mellitus Type 2 at the Regional Public Hospital of Pasuruan with respondents who counted 96 patients. The purpose of using PSS-SVM, in this case, is to reduce the confounding effects (sports activity) that influence disease complications. In strata of 4 reduce the largest bias with the percent bias reduction (PBR) is 86.39% with the smallest standard error is 0.103.

**Keywords:** accuration, diabetes mellitus, PBR, propensity score stratification, SVM.

---

\* Received Mar 2020; Accepted Nov 2020; Published online on Nov 2020

## 1. Pendahuluan

Dalam bidang kesehatan, penelitian *Randomized Controlled Trials* (RCT) tidak selalu dapat dilakukan karena berhubungan langsung dengan nyawa manusia sehingga menggunakan penelitian observasional. Namun, penelitian observasional menyebabkan peneliti tidak dapat melakukan kontrol atas perlakuan, sehingga kelompok perlakuan dan kontrol sering tidak seimbang. Hal ini menyebabkan perbedaan besar pada kovariat yang diamati sehingga menyebabkan perkiraan efek perlakuan menjadi bias (D'Agostino, 1998). Kovariat ini merupakan variabel pengendali, sehingga harus memiliki pengaruh yang seimbang, baik untuk variabel terikat pada kelompok perlakuan maupun kelompok kontrol. Salah satu metode alternatif yang semakin sering digunakan dalam analisis statistika bidang kesehatan adalah *Propensity Score* (PS). Jika skor kecenderungan tersebut dibagi ke dalam kelompok strata-strata yang memiliki skor kecenderungan serupa, maka dikenal dengan *Propensity Score Stratification* (PSS-SVM). Tujuan PSS-SVM untuk menyeimbangkan variabel yang diamati antara kelompok perlakuan dan kelompok kontrol dengan membagi menjadi beberapa kelompok strata sehingga didapatkan kecenderungan pengamatan yang seimbang atau dalam *propensity score* disebut dengan istilah *balance*. PSS dilakukan agar estimasi efek perlakuan dalam stratum akan sama dan mengurangi efek *confounding* atau variabel perancu dalam penelitian (McCandless et al., 2009). Pada penelitian sebelumnya menyebutkan bahwa *propensity score stratification* menggunakan regresi logistik dapat mengurangi bias akibat *confounding* (Merdeka, 2017). Namun, beberapa penelitian yang sudah ada, masih menggunakan metode statistika klasik dan hanya memenuhi beberapa kondisi tertentu. Oleh karena itu, pada penelitian ini menggunakan pendekatan metode *machine learning* dalam estimasi *propensity score* dengan asumsi yang lebih sedikit atau akurasi yang lebih besar (Westreich et al., 2010). *Propensity score SVM* (PS-SVM) yang sebelumnya dilakukan menghasilkan bahwa PS-SVM prediksi lebih baik, yaitu mampu memprediksi antigenic epitopes dengan sensitivitas sebesar (80,1%) dan nilai AUC sebesar 0,702 (Zhao et al., 2006).

Adapun pada penelitian ini, penerapan metode *Propensity Score Stratification – Support Vector Machine* (PSS-SVM) diaplikasikan pada studi kasus data pasien penderita Diabetes Melitus. Penyakit Diabetes Melitus merupakan penyebab kematian peringkat ke-3 tertinggi di Indonesia pada tahun 2012, dengan persentase korban meninggal sebesar 6,5 % dari 328,5 ribu jiwa ([WHO], 2015). Penyakit Diabetes Melitus ditandai dengan tingginya kadar glukosa darah yaitu melebihi 200 mg/dL, dan keadaan hiperglikemia ini bersifat menahun sehingga mempengaruhi seluruh sistem tubuh (Misnadiarly, 2006). Salah satu klasifikasi dari penyakit Diabetes Melitus yaitu DM Tipe 2 (*non-insulin-dependent* atau *adult-onset diabetes*) yaitu ketidakmampuan tubuh menggunakan insulin secara efektif, sehingga berat badan penderita menjadi berlebih dan susah melakukan aktivitas fisik ([KEMENKES RI], 2013). Selain itu, faktor yang mempengaruhi DM tipe 2 antara lain riwayat keturunan, obesitas, jenis kelamin dan faktor lingkungan seperti diet, obesitas dan aktivitas olahraga (Gungor et al., 2005). Faktor risiko yang mempengaruhi Diabetes Melitus tersebut saling mempengaruhi dalam menyebabkan kondisi yang lebih parah, misalnya adanya komplikasi penyakit. Komplikasi penyakit pada kasus Diabetes

Melitus Tipe 2 dipengaruhi oleh faktor-faktor penyebab yang saling berkorelasi atau berhubungan dan menyebabkan adanya variabel perancu atau *confounding* pada variabel prediktor. Oleh karena itu, dengan menggunakan metode *propensity score* stratifikasi SVM (PSS-SVM) dapat mereduksi efek *confounding* tersebut. Penggunaan PSS SVM diharapkan dapat mereduksi adanya efek *confounding* dan diperoleh *balance* kovariat. Hal ini dapat menghilangkan pengaruh dari *confounding* terhadap kelompok perlakuan dan kelompok kontrol untuk memperoleh estimasi efek perlakuan yang tidak bias. Sehingga pada penelitian ini, peneliti melakukan analisis “Penggunaan Propensity Score-Support Vector Machine untuk Mengestimasi Efek Perlakuan Aktivitas Olahraga pada Penderita Diabetes Melitus Tipe 2”, dengan menggunakan data pasien kasus komplikasi penyakit Diabetes Melitus Tipe 2 di RSUD Kabupaten Pasuruan.

## 2. Metodologi

### 2.1 Bahan dan Data

Penelitian ini menggunakan data sekunder dari survei dan rekam medis pasien Diabetes Melitus Tipe 2 pada Maret 2016 di Rumah Sakit Umum Daerah (RSUD), Kabupaten Pasuruan, Jawa Timur dengan jumlah responden sebanyak 96 pasien. Variabel respon yang digunakan dalam penelitian ini adalah ada tidaknya komplikasi pada pasien Diabetes Melitus (Y) terdiri dari dua kategori yaitu ada komplikasi dan tidak ada komplikasi, sedangkan terdapat tujuh variabel prediktor (X). Adapun variabel penelitian dalam penelitian ditunjukkan pada Tabel 1 berikut.

Tabel 1: Variabel Penelitian.

Variabel	Keterangan	Satuan/Kategori
Y	Komplikasi Penyakit	1. Ada (0) 2. Tidak ada (1)
X <sub>1</sub>	Usia	Tahun
X <sub>2</sub>	Jenis Kelamin	1. Perempuan (0) 2. Laki-laki (1)
X <sub>3</sub>	Riwayat Genetik	1. Ada (0) 2. Tidak ada (1)
X <sub>4</sub>	Obesitas	1. Ya , BMI $\geq 27,5$ (0) 2. Tidak , BMI $\leq 27,5$ (1)
X <sub>5</sub>	Hipertensi	1. Hipertensi jika tekanan darah sistolik kurang dari 130 mmHg (0) 2. Tidak hipertensi jika tekanan darah sistolik sama dengan atau lebih dari 130 mmHg (1)
X <sub>6</sub>	Lama Menderita	Tahun
X <sub>7</sub>	Aktivitas Olahraga	1. Kurang (0) 2. Cukup (1)

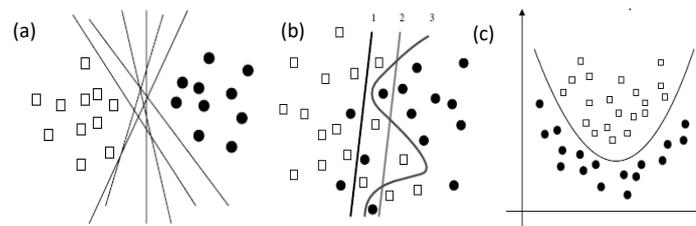
### 2.2 Metode Penelitian

#### a. Support Vector Machine (SVM)

Gambar 1 menunjukkan ilustrasi untuk pembagian SVM klasifikasi. SVM untuk klasifikasi dapat dibedakan menjadi klasifikasi linier, baik *linearly separable* maupun *linearly non-separable* serta klasifikasi non linear. Prinsip dasar SVM adalah *linear*

*classifier*, tetapi pada kenyataan atau kasus data *real*, tidak semua data bersifat linier sehingga untuk mencari bidang pemisah secara linier sulit dilakukan. Oleh karena itu, untuk mengklasifikasikan data yang tidak linear dapat diselesaikan menggunakan SVM non-linear dengan menambahkan fungsi *kernel trick* (Härdle et al., 2014). Adapun nilai dari fungsi klasifikasi (*score*) dapat dihitung menggunakan persamaan (1) berikut.

$$f(x) = \sum_{i=1}^n \hat{\alpha}_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b \tag{1}$$



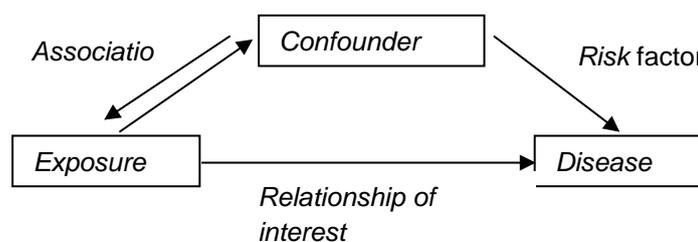
Gambar 1: a) *Linearly separable*; (b) *Linearly Non separable*; (c) *Nonlinear* (Härdle et al., 2014).

Beberapa fungsi kernel yang umum digunakan antara lain *Kernel Linear*, *Polynomial Kernel*, *Radial Basis Function* dan *Sigmoid Kernel* (Min & Lee, 2005). Selanjutnya satu diantara fungsi kernel tersebut dapat membentuk fungsi keputusan pada persamaan (2).

$$\hat{g}(x) = \text{sign} \left( \sum_{i=1}^n \hat{\alpha}_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + \hat{b} \right) \tag{2}$$

**b. Faktor Confounding**

*Confounder* dapat didefinisikan sebagai : (a) faktor yang berhubungan dengan variabel bebas lainnya (*exposure variable*) (b) faktor risiko yang independen terhadap *outcome*. Jadi, *confounder* tidak seharusnya secara langsung dihubungkan antara variabel *exposure* dan *outcome* (Winkelmayer & Kurth, 2004). Adapun untuk menjelaskan fenomena *confounding*, perlu diperhatikan hubungan antara *exposure* dan terjadinya suatu penyakit pada Gambar 2 berikut.



Gambar 2: Ilustrasi Faktor *Confounding*.

Variabel *confounding* dapat ditentukan secara teori dan secara empiris melalui pengujian. Secara empiris, uji independensi dapat menggunakan uji *chi-square* untuk memeriksa hubungan antara dua variabel (Agresti, 2003).

### c. Propensity Score Stratification (PSS) Menggunakan Support Vector Machine (SVM)

*Propensity score* (PS) didefinisikan sebagai probabilitas bersyarat, dengan nilai antara 0 dan 1, dengan subjek atau pengamatan diberikan perlakuan berdasarkan kelompok kovariat yang diamati (Rosenbaum & Rubin, 1983). Berikut bentuk persamaan matematis *propensity score*.

$$e(\mathbf{x}_i) = P(Z_i = 1 | \mathbf{X}_i = \mathbf{x}_i) = \frac{\exp(f(\mathbf{x}_i))}{1 + \exp(f(\mathbf{x}_i))} \quad (3)$$

Pada penelitian ini estimasi *propensity score* menggunakan salah satu metode *machine learning* yaitu *support vector machine* (SVM). Prinsip PS SVM yaitu dengan mensubstitusi model SVM pada persamaan (1) ke dalam model *propensity score* pada persamaan (3), sehingga diperoleh persamaan (4) berikut.

$$e(\mathbf{x}_i) = \frac{\exp(f(\mathbf{x}_i))}{1 + \exp(f(\mathbf{x}_i))} = \frac{\exp(\hat{z}_i)}{1 + \exp(\hat{z}_i)} \quad (4)$$

*Propensity Score Stratification* (PSS) merupakan salah satu metode *propensity score* untuk mengontrol perbedaan sistematis antara kelompok perlakuan dan kelompok kontrol yaitu dengan melakukan pengelompokan subjek atau pengamatan ke dalam strata-strata yang sama berdasarkan nilai estimasi *propensity score*. Tujuan dari PSS ini adalah untuk mendapatkan kelompok strata yang menghasilkan variabel yang seimbang (*balance*) antara kelompok perlakuan dan kelompok kontrol.

### d. Uji Balance Kovariat

Pada *propensity score*, pengaruh dari kovariat terhadap kelompok perlakuan dan kelompok kontrol dapat dihilangkan dengan cara menyeimbangkan kovariat (*balance* kovariat), sehingga diperoleh estimasi efek perlakuan yang tidak bias. Keseimbangan kovariat antar kelompok perlakuan dan kontrol dapat diketahui salah satunya dengan membandingkan variansi statistik seperti rata-rata baik dengan uji kesamaan dua rata-rata (uji-t), atau uji kesamaan dua proporsi pada kovariat (Li, 2013). Jika bentuk data kontinyu, maka menggunakan uji-t, sedangkan jika data berbentuk kategori menggunakan uji-z untuk menguji perbedaan dua kelompok.

Hipotesis untuk menguji perbedaan proporsi pada data kategori sebagai berikut.

$H_0 : p_t = p_c$  (Proporsi antara kelompok perlakuan dan kelompok kontrol adalah sama atau tidak ada perbedaan proporsi antara kelompok perlakuan dan kontrol)

$H_1 : p_t \neq p_c$  (Proporsi antara kelompok perlakuan dan kelompok kontrol tidak sama atau ada perbedaan proporsi antara kelompok perlakuan dan kontrol)

Statistik uji yang digunakan sesuai persamaan (5).

$$Z = \frac{\hat{p}_t - \hat{p}_c}{\sqrt{\hat{p}(1-\hat{p}) \left[ \frac{1}{n_t} + \frac{1}{n_c} \right]}} \tag{5}$$

dengan  $\hat{p} = \frac{x_t + x_c}{n_t + n_c}$ ;  $\hat{p}_t = \frac{x_t}{n_t}$ ;  $\hat{p}_c = \frac{x_c}{n_c}$

dimana  $\hat{p}$  adalah proporsi gabungan,  $n_t$  adalah banyak observasi kelompok perlakuan,  $n_c$  adalah banyak observasi kelompok kontrol,  $\hat{p}_t$  adalah proporsi kelompok perlakuan, dan  $\hat{p}_c$  adalah proporsi kelompok kontrol. Daerah kritis untuk taraf signifikansi  $\alpha$  adalah tolak  $H_0$  jika  $|Z_{hit}| > Z_{\alpha/2}$  atau  $p\text{-value} < \alpha$ .

Langkah selanjutnya untuk menguji keseimbangan kovariat data kontinyu menggunakan uji-t. Adapun hipotesis untuk pengujian kesamaan variansi dengan statistik uji pada persamaan (6) sebagai berikut,

$H_0 : \mu_t = \mu_c$  (rata-rata kelompok perlakuan dan kelompok kontrol sama atau tidak ada perbedaan rata-rata antara kelompok perlakuan dan kontrol)

$H_1 : \mu_t \neq \mu_c$  (rata-rata kelompok perlakuan dan kelompok kontrol tidak sama atau ada perbedaan rata-rata antara kelompok perlakuan dan kontrol)

$$t = \frac{\bar{x}_t - \bar{x}_c}{\sqrt{\frac{s_t^2}{n_t} + \frac{s_c^2}{n_c}}} \tag{6}$$

dengan  $df = \frac{\left[ \left( \frac{s_t^2}{n_t} \right) + \left( \frac{s_c^2}{n_c} \right) \right]^2}{\left[ \left( \frac{s_t^2}{n_t} \right) / (n_t - 1) \right] + \left[ \left( \frac{s_c^2}{n_c} \right) / (n_c - 1) \right]}$

$\bar{x}_t$  merupakan rata-rata dari kelompok perlakuan,  $\bar{x}_c$  merupakan rata-rata kelompok kontrol,  $n_t$  merupakan banyaknya observasi pada kelompok perlakuan,  $n_c$  merupakan banyaknya observasi pada kelompok kontrol,  $s_t^2$  merupakan varians kelompok perlakuan dan  $s_c^2$  merupakan varians kelompok kontrol, dan  $s$  merupakan simpangan baku gabungan. Tolak  $H_0$  jika  $|t_{hit}| > t_{\frac{\alpha}{2}, df}$  atau  $p\text{-value} < \alpha$ .

**e. Ketepatan Klasifikasi**

Performasi klasifikasi dapat ditentukan menggunakan nilai yang terdapat dalam *confusion matrix*, yaitu matriks yang berisi informasi tentang klasifikasi aktual dan prediksi yang dilakukan oleh sistem klasifikasi (Provost, 1998).

Tabel 2: Confussion Matrix

Actual	Prediction	
	I(Negative)	II(Positive)
Negative	True Negative (TN)	False Positive (FP)
Positive	False Negative (FN)	True Positive (TP)

$$\text{Akurasi} = \frac{(\text{TN} + \text{TP})}{(\text{TN} + \text{TP} + \text{FN} + \text{FP})} \quad (7)$$

Jika setiap strata ke- $i$  mempunyai akurasi tertentu, maka akurasi pada keseluruhan strata dapat dihitung dengan persamaan (8).

$$\text{Acc} = \sum_{i=1}^c (\text{Acc}_i \times P_i) \quad (8)$$

dengan  $\text{Acc}_i$  menunjukkan akurasi strata ke- $i$  dan  $P_i$  merupakan persentase atau proporsi sampel dalam strata ke- $i$  (Yu et al., 2013).

#### f. Estimasi Average Treatment Effect (ATE)

Secara ringkas tahapan untuk estimasi *Average Treatment Effect* (ATE) dengan menggunakan *propensity score* antara lain : membagi subjek atau melakukan stratifikasi ke dalam  $K$  subkelas yang homogen, melakukan estimasi Average Treatment Effect (ATE)  $\hat{\delta}$  sesuai persamaan (9), dan menghitung standar error dari ATE dengan persamaan (10).

$$\hat{\delta} = \sum_{k=1}^K \frac{n_{tk} + n_{ck}}{n_t + n_c} (\bar{Y}_{tk} - \bar{Y}_{ck}) \quad (9)$$

$$SE(\hat{\delta}) = \sqrt{\sum_{k=1}^K \left( \frac{n_{tk} + n_{ck}}{n_t + n_c} \right) \left( \frac{s_{tk}^2}{n_{tk}} + \frac{s_{ck}^2}{n_{ck}} \right)} \quad (10)$$

dengan  $n_t$  adalah banyak subjek kelompok perlakuan,  $n_c$  adalah banyak subjek kelompok kontrol,  $n_{tk}$  dan  $n_{ck}$  adalah banyak subjek kelompok perlakuan strata ke- $k$ ,  $n_{tk}$  adalah banyak subjek kelompok kontrol strata ke- $k$ ,  $s_{tk}^2$  dan  $s_{ck}^2$  masing-masing adalah varians respon untuk kelompok perlakuan strata ke- $k$  dan kelompok kontrol strata ke- $k$ ,  $k$  adalah banyaknya kelompok (Tu & Zhou, 2002). Selanjutnya diperoleh statistik uji berdasarkan persamaan (9) dan (10) untuk estimasi sebagai berikut.

$$Z = \frac{\hat{\delta}}{SE(\hat{\delta})} \quad (11)$$

dengan  $\hat{\delta}$  merupakan rata-rata efek perlakuan (ATE), dan  $SE(\hat{\delta})$  merupakan standar error dari ATE. Daerah kritis untuk taraf signifikansi  $\alpha$  adalah tolak  $H_0$  jika  $|Z_{\text{hit}}| > Z_{\alpha/2}$  atau  $p\text{-value} < \alpha$ .

### g. Evaluasi Propensity Score Stratification

Salah satu kriteria kebaikan metode PSS dapat dilihat dari seberapa besar bias yang mampu direduksi untuk setiap kovariat maupun bias secara keseluruhan. Adapun untuk menghitung besarnya bias yang direduksi untuk masing-masing kovariat yaitu dengan formula sebagai berikut (Pan & Bai, 2015).

$$PBR = \frac{B_{\text{BeforeStrata}} - B_{\text{AfterStrata}}}{B_{\text{BeforeStrata}}} \times 100\% \quad (12)$$

dan

$$B = p_1(x_p) - p_0(x_0) \quad (13)$$

dengan PBR menunjukkan *Percent Bias Reduction*, B merupakan selisih rata-rata dari kelompok perlakuan dan kelompok kontrol untuk setiap kovariat,  $p_1(x_p)$  dan  $p_0(x_0)$  masing-masing merupakan proporsi kovariat ke-p untuk kelompok perlakuan dan kelompok kontrol.  $B_{\text{BeforeStrata}}$  merupakan selisih rata-rata kelompok perlakuan dan kontrol sebelum dilakukan stratifikasi, sedangkan  $B_{\text{AfterStrata}}$  merupakan selisih rata-rata kelompok perlakuan dan kontrol setelah dilakukan stratifikasi.

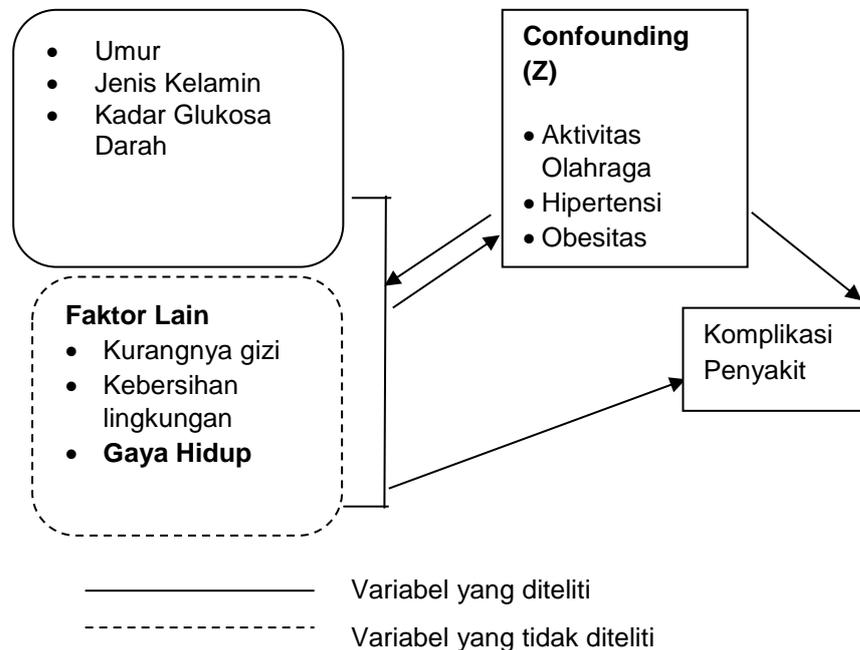
### h. Diabetes Melitus Tipe 2 dan Faktor yang Mempengaruhi Komplikasi Penyakit Penderita DM Tipe 2

Penyakit Diabetes Melitus ditandai dengan tingginya kadar glukosa darah yaitu melebihi 200 mg/dL, dan keadaan hiperglikemia ini bersifat menahun sehingga mempengaruhi seluruh sistem tubuh (Misnadiarly, 2006). Salah satu jenis dari penyakit Diabetes Melitus adalah DM tipe 2 (atau dikenal dengan non-insulin-dependent), disebabkan ketidakmampuan tubuh menggunakan insulin secara efektif. Penderita DM tipe 2 biasanya memiliki berat badan berlebih sehingga menghambat aktivitas fisik ([KEMENKES RI], 2013).

Penderita DM mempunyai faktor risiko yang berbeda-beda. Beberapa faktor risiko DM antara lain terdiri dari faktor usia, faktor keturunan (genetik), faktor obesitas, faktor penyakit penyerta dan faktor demografi meliputi peningkatan jumlah penduduk, kurangnya gizi dan keadaan lingkungan (Sustrani et al., 2006). Bagi penderita DM, juga meningkatkan risiko terkena hipertensi karena gangguan insulin pada tubuh dapat menyebabkan peningkatan tekanan darah. Selain hipertensi, obesitas juga menjadi salah satu faktor risiko penyebab terjadinya penyakit DM. Obesitas dapat dikurangi antara lain dengan menjaga pola makan, melakukan aktivitas fisik atau olahraga dan sebagainya. Peran aktivitas fisik pada penderita DM sangat penting. Hal ini sesuai dengan hasil penelitian yang dilakukan oleh (Sharoh & Salmiyati, 2017), (Afriza, 2015), dan (Sinaga & Hondro, 2011) bahwa adanya efektivitas aktivitas fisik dapat menurunkan kadar glukosa darah pada penderita DM. Salah satu contoh aktivitas fisik adalah melakukan senam.

Berdasarkan uraian di atas, beberapa variabel antara lain : aktivitas olahraga (senam), hipertensi dan obesitas berpotensi menjadi variabel *confounding*. Selanjutnya secara empiris dilakukan uji dependensi untuk menentukan variabel

*confounding*. Gambaran kerangka konsep dalam penelitian ini dapat ditunjukkan pada Gambar 3 berikut.



Gambar 3: Kerangka Konsep Faktor Resiko Komplikasi Penyakit pada Diabetes Melitus Tipe 2.

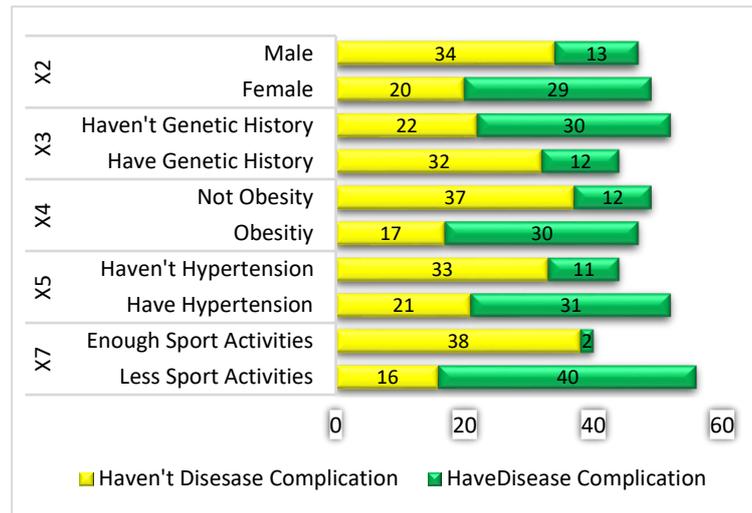
Adapun langkah analisis *Propensity Score Stratification* dengan *Support Vector Machine (PSS)-SVM* pada kasus Diabetes Melitus Tipe 2 sebagai berikut.

1. Mendeskripsikan secara deskriptif karakteristik pasien.
2. Menguji hubungan antar variabel untuk menentukan variabel yang menjadi variabel *confounding (Z)*.
3. Menghitung *estimasi PSS-SVM*. Adapun fungsi kernel yang digunakan yaitu *radial basis function* dan pemilihan parameter optimal pada SVM menggunakan metode taguchi dengan rentang nilai *cost* ( $C=1, 10, 50, 100$ ) dan nilai *gamma* ( $\gamma = 0,1; 1; 3; 5; 10; 12$ ) yang telah ditentukan.
4. Berdasarkan nilai *propensity score* yang diurutkan dari terkecil ke terbesar, subjek dibagi ke dalam 2, 3, 4, dan 5 strata.
5. Pada masing-masing strata, dilakukan uji asumsi *balance* pada kelompok perlakuan dan kelompok kontrol.
6. Menghitung akurasi klasifikasi pada setiap strata.
7. Menghitung nilai estimasi *average treatment effect (ATE)* dan pengujian signifikansi pada estimasi ATE
8. Selanjutnya, strata yang memenuhi kriteria *balance kovariat* dihitung nilai *Percent Bias Reduction (PBR)*.

### 3. Hasil dan Pembahasan

#### 3.1 Gambaran Umum Karakteristik Pasien Diabetes Melitus

Karakteristik pasien dari 96 pasien Diabetes Melitus dapat dilihat melalui analisis statistika deskriptif. Gambaran umum untuk variabel jenis kelamin, riwayat genetik, obesitas, hipertensi, dan aktivitas olahraga pada pasien Diabetes Melitus ditunjukkan pada Gambar 4.



Gambar 4: Karakteriksi Pasien DM Tipe 2.

Dilihat dari jenis kelamin, dari 96 responden, pasien DM dengan jenis kelamin laki-laki yang mempunyai penyakit komplikasi lebih sedikit dibanding pasien dengan jenis kelamin perempuan yang mempunyai komplikasi penyakit, yaitu hanya 13 pasien laki-laki. Pada penelitian ini, meskipun pasien mempunyai riwayat genetik, tidak selalu mengalami komplikasi penyakit, hal ini dapat dilihat bahwa sepertiga dari responden atau 32 pasien yang mempunyai riwayat genetik tetapi tidak mengalami komplikasi penyakit. Pasien DM yang mengalami komplikasi dan obesitas sebesar 31%. Pasien yang tidak hipertensi cenderung tidak mengalami komplikasi penyakit yaitu sebanyak 33 pasien, dan sebaliknya, pasien yang hipertensi akan cenderung mengalami komplikasi penyakit sebesar 31 pasien.

Ada atau tidak suatu komplikasi penyakit pada pasien Diabetes Melitus dapat dikurangi atau diminimalkan dengan melakukan suatu aktivitas olahraga. Pasien dengan aktivitas olahraga cukup lebih banyak untuk tidak ada komplikasi yaitu sebanyak 38 pasien, dan sebaliknya pasien dengan aktivitas olahraga kurang lebih banyak untuk mengalami komplikasi penyakit sebanyak 40 pasien. Oleh karena itu, diduga aktivitas olahraga dapat meminimalkan adanya komplikasi penyakit pada pasien Diabetes Melitus.

#### 3.2 Analisis Propensity Score Stratification

##### a. Pemilihan Variabel Confounding

Pada tinjauan pustaka yang telah dijelaskan di atas bahwa variabel yang memengaruhi adanya ada atau tidak komplikasi penyakit pasien Diabetes Melitus Tipe 2 antara lain jenis kelamin, usia, riwayat genetik, status obesitas, status hipertensi,

lama menderita dan aktivitas olahraga. Dari variabel yang berpengaruh tersebut terdapat beberapa variabel yang diduga sebagai *confounding*, selanjutnya dilakukan pengujian antara variabel yang diduga sebagai *confounding* dengan kejadian infeksi oportunistik (Y) ditampilkan pada Tabel 3.

Tabel 3: Pegujian Hubungan antara Variabel X dan Y.

Variabel	$\chi^2$	df	p-value	Keputusan
X <sub>2</sub> vs Y	9,69	1	0,002*	Tolak H <sub>0</sub>
X <sub>3</sub> vs Y	8,97	1	0,003*	Tolak H <sub>0</sub>
X <sub>4</sub> vs Y	15,09	1	0,000*	Tolak H <sub>0</sub>
X <sub>5</sub> vs Y	11,60	1	0,001*	Tolak H <sub>0</sub>
X <sub>7</sub> vs Y	41,84	1	0,000*	Tolak H <sub>0</sub>

\*Significancant at level ( $\alpha = 0,05$ )

Berdasarkan Tabel 3, diketahui bahwa terdapat hubungan antara X<sub>2</sub>, X<sub>3</sub>, X<sub>4</sub>, X<sub>5</sub> dan X<sub>7</sub> memiliki hubungan yang signifikan dengan variabel komplikasi penyakit. Berdasarkan beberapa penelitian sebelumnya yang menyebutkan bahwa aktivitas olahraga dapat mempengaruhi ada atau tidak komplikasi penyakit pada pasien DM Tipe 2, maka selanjutnya dengan melalui pengujian empiris untuk melihat hubungan tiap variabel dan variabel yang berpotensi menjadi *confounding* menggunakan statistik uji *chi-square* yang ditampilkan pada Tabel 4.

Tabel 4: Pengujian Hubungan Antar Kovariat.

Variabel	$\chi^2$	df	p-value	Keputusan
X <sub>3</sub> vs X <sub>2</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>4</sub> vs X <sub>2</sub>	4,19	1	0,041*	Tolak H <sub>0</sub>
X <sub>5</sub> vs X <sub>2</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>7</sub> vs X <sub>2</sub>	7,06	1	0,008*	Tolak H <sub>0</sub>
X <sub>2</sub> vs X <sub>3</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>4</sub> vs X <sub>3</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>5</sub> vs X <sub>3</sub>	2,48	1	0,115	Gagal Tolak H <sub>0</sub>
X <sub>7</sub> vs X <sub>3</sub>	10,15	1	0,001*	Tolak H <sub>0</sub>
X <sub>2</sub> vs X <sub>4</sub>	4,19	1	0,004*	Tolak H <sub>0</sub>
X <sub>3</sub> vs X <sub>4</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>5</sub> vs X <sub>4</sub>	3,46	1	0,063	Gagal Tolak H <sub>0</sub>
X <sub>7</sub> vs X <sub>4</sub>	7,43	1	0,006*	Tolak H <sub>0</sub>
X <sub>2</sub> vs X <sub>5</sub>	0,39	1	0,528	Gagal Tolak H <sub>0</sub>
X <sub>3</sub> vs X <sub>5</sub>	2,48	1	0,115	Gagal Tolak H <sub>0</sub>
X <sub>4</sub> vs X <sub>5</sub>	3,46	1	0,063	Gagal Tolak H <sub>0</sub>
X <sub>7</sub> vs X <sub>5</sub>	11,60	1	0,000*	Tolak H <sub>0</sub>
X <sub>2</sub> vs X <sub>7</sub>	7,06	1	0,007*	Tolak H <sub>0</sub>
X <sub>3</sub> vs X <sub>7</sub>	10,15	1	0,001*	Tolak H <sub>0</sub>
X <sub>4</sub> vs X <sub>7</sub>	7,43	1	0,006*	Tolak H <sub>0</sub>
X <sub>5</sub> vs X <sub>7</sub>	5,54	1	0,019*	Tolak H <sub>0</sub>

\*Significancant at level ( $\alpha = 0,05$ )

Statistik uji *chi-square* pada taraf signifikansi ( $\alpha$ ) sebesar 0,05 pada hasil pengujian pada Tabel 4 dapat disimpulkan bahwa pada variabel aktivitas olahraga (X<sub>7</sub>) memiliki empat variabel yang saling berhubungan yaitu variabel X<sub>2</sub>, X<sub>3</sub>, X<sub>4</sub> dan X<sub>5</sub>. Apabila dibandingkan dengan variabel lainnya maka variabel aktivitas olahraga (X<sub>7</sub>) berpotensi

sebagai *confounding* karena variabel tersebut memiliki hubungan signifikan paling banyak dengan variabel lain. Oleh karena itu variabel aktivitas olahraga ( $X_7$ ) dipilih sebagai *confounding*.

#### b. Estimasi Nilai Propensity Score (SVM)

Pada penelitian ini, estimasi *propensity score* menggunakan SVM. Sebelum mendapatkan nilai *propensity score*, maka perlu didapatkan fungsi klasifikasi SVM dalam bentuk *score* terlebih dahulu dengan fungsi seperti pada persamaan (1). Menggunakan metode taguchi dan fungsi kernel *radial basis function* diperoleh parameter optimum untuk  $C=1$  dan  $\gamma=0,1$ . Tabel 5 berikut menunjukkan fungsi klasifikasi dalam bentuk *score* menggunakan fungsi SVM.

Tabel 5: Fungsi Klasifikasi (Score).

No	$f(\mathbf{x}_i)$						
1-7	0,098	0,058	0,048	0,064	0,069	0,209	0,044
8-14	0,063	0,286	0,049	0,946	0,724	0,068	0,049
:	:	:	:	:	:	:	:
92 -96	0,497	0,764	0,049	0,049	0,615		

Langkah selanjutnya yaitu mengestimasi *propensity score* untuk 96 responden dengan menggunakan *score* pada Tabel 5 sesuai persamaan (4) ditunjukkan pada Tabel 6 berikut.

Tabel 6: Nilai Propensity Score.

No	$e(\mathbf{x}_i)$						
1-7	0,524	0,514	0,512	0,516	0,517	0,552	0,511
8-14	0,516	0,571	0,512	0,720	0,674	0,517	0,512
:	:	:	:	:	:	:	:
92-96	0,622	0,682	0,512	0,512	0,649		

Berdasarkan Tabel 6 diperoleh bahwa nilai *propensity score* untuk masing-masing responden dengan nilai *propensity score* minimum adalah 0,511 dan nilai *maximum* adalah 0,720.

#### c. Pembagian ke dalam Strata dan Mengevaluasi Balance Kovariat

Nilai *propensity score* yang diperoleh pada langkah sebelumnya, kemudian dilakukan pengurutan mulai dari nilai paling besar sampai nilai paling kecil, dan dilakukan pembagian ke dalam strata-strata. Pembagian subjek ke dalam strata untuk memperoleh kelompok strata dengan kovariat yang seimbang atau *balance*. Keseimbangan ini bertujuan untuk menghilangkan pengaruh kovariat terhadap kelompok perlakuan dan kelompok kontrol sehingga tidak ada perbedaan antara kelompok perlakuan dan kelompok kontrol. Adapun hasil pengujian dari nilai p-value untuk uji z dan uji t masing-masing variabel dapat dilihat pada Tabel 7 berikut.

Tabel 7: Nilai p-value uji z dan uji t pada Pengujian Balance Kovariat.

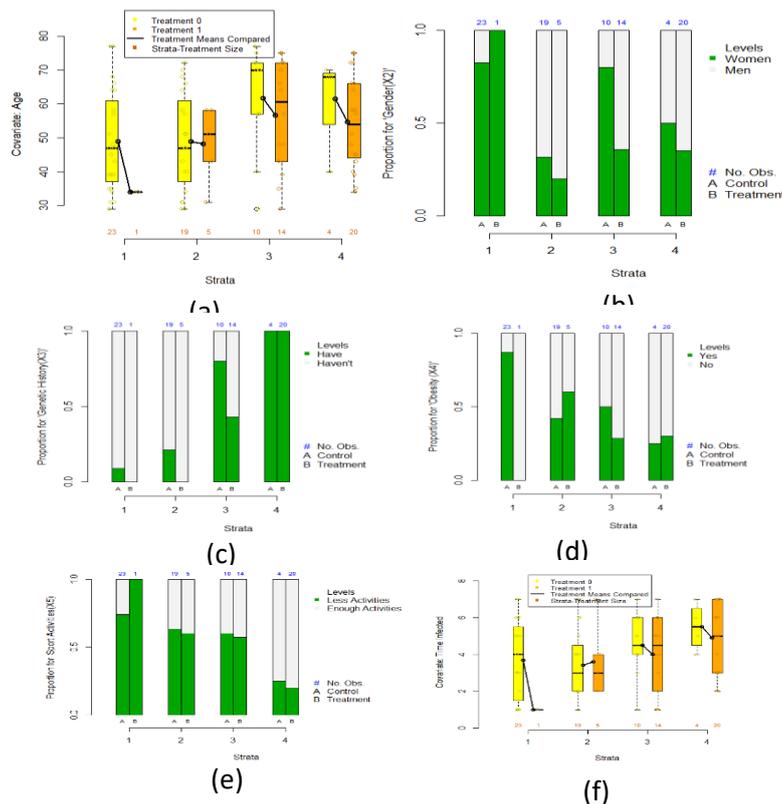
Banyak Strata	Strata ke – k	<i>p-value</i>					
		X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Sebelum	Strata	0,518	<b>0,008</b>	<b>0,001</b>	<b>0,019</b>	<b>0,006</b>	0,283
2	1	0,585	0,226	0,322	0,425	0,906	0,677
	2	0,204	<b>0,022</b>	0,474	0,369	0,343	0,645
3	1	0,952	0,119	0,492	0,614	0,164	0,480
	2	0,650	0,348	0,299	0,409	0,722	1,000
	3	0,762	0,533	NaN	0,399	0,497	0,506
4	1	0,952	0,648	0,758	<b>0,022</b>	0,555	0,480
	2	0,919	0,612	0,261	0,475	0,897	0,868
	3	0,455	<b>0,032</b>	0,069	0,285	0,889	0,526
	4	0,431	0,572	NaN	0,841	0,822	0,431
5	1	0,063	NaN	NaN	NaN	NaN	0,992
	2	0,728	0,435	0,329	0,435	0,906	0,662
	3	0,269	0,714	0,163	0,292	0,110	0,577
	4	0,919	0,147	0,130	0,629	0,279	0,919
	5	<b>0,002</b>	0,372	NaN	0,440	0,288	<b>0,002</b>

Ket : ck adalah control kelas ke-k, tk adalah treatment kelas ke-k  
 Taraf signifikansi ( $\alpha$ ) yang digunakan 0,05.

Pengujian *balance* antara kelompok perlakuan dan kelompok kontrol untuk setiap kovariat pada masing-masing strata dapat dilihat dari nilai p-value pada *Tabel 7*. Sebelum dilakukan stratifikasi, ada empat variabel yang memberikan nilai p-value  $<(\alpha)$  yaitu variabel jenis kelamin ( $X_2$ ), variabel riwayat genetik ( $X_3$ ), dan variabel status hipertensi ( $X_5$ ). Hal ini menunjukkan bahwa terdapat perbedaan proporsi antara kelompok perlakuan dan kontrol sebelum stratifikasi pada kovariat  $X_2$ ,  $X_3$ , dan  $X_5$  atau dapat dikatakan belum *balance*. Kemudian dilakukan stratifikasi pada masing-masing kovariat dengan pembagian strata sebanyak 2, 3, 4, dan 5. Pada strata sebanyak 2,3, dan 4 diketahui bahwa kelompok perlakuan dan kelompok kontrol untuk semua kovariat telah *balance* yang dibuktikan dengan nilai *p-value* pada masing-masing strata lebih dari 0,05. Akan tetapi, pada strata sebanyak 5 diketahui ada kovariat yang belum *balance* yang ditunjukkan dengan nilai p-value  $< 0,05$  yaitu kovariat  $X_1$  dan  $X_6$  pada strata ke-5. Selain itu, pada strata sebanyak 5 terdapat beberapa nilai NaN (*Not a Number*) yang disebabkan karena terdapat satu kelompok perlakuan atau kelompok kontrol yang tidak berisi anggota pengamatan, sehingga nilai statistik uji

atau *p-value* tidak dapat dihitung atau tidak dapat ditampilkan.

Gambar 5 adalah salah satu grafik untuk mendukung hasil uji balance secara visual untuk masing-masing kovariat pada strata sebanyak 4. Dapat dilihat bahwa proporsi antara kelompok perlakuan dan kelompok kontrol tidak jauh berbeda di setiap strata untuk masing-masing kovariat. Secara keseluruhan, tidak terdapat perbedaan proporsi antara kelompok perlakuan dan kelompok kontrol pada semua strata yang telah dilakukan.



Gambar 5: Balance Kovariat pada Masing-Masing Kovariat untuk Strata Sebanyak 4 (a) Usia (b) Pengetahuan, (c) Sikap, (d) Konsep Diri, (e) Dukungan Keluarga dan (f) Lama Menderita.

**d. Tingkat Akurasi Klasifikasi PSS-SVM**

Pada langkah sebelumnya telah dilakukan pembagian nilai PSS-SVM ke dalam strata 2, 3, 4, dan 5. Selanjutnya dari setiap strata dilakukan klasifikasi berdasarkan nilai *propensity score* dan didapatkan akurasi dari PSS-SVM yang ditampilkan pada Tabel 8 berikut.

Tabel 8 menunjukkan tingkat akurasi masing-masing strata dan per bagian strata pada strata tertentu. Seperti pada penjelasan pada Tabel 7 bahwa pada strata sebanyak 5 terdapat beberapa subjek yang tidak tersedia sehingga menyebabkan akurasi pada strata sebanyak 5 tidak dapat dilakukan perhitungan. Jika dilihat dari tingkat akurasi yang dihasilkan pada bagian strata, maka pada strata sebanyak 4,

bagian strata ke-5 memiliki tingkat akurasi tertinggi yaitu sebesar 0,917 atau 91,7 persen. Apabila dilihat tingkat akurasi secara keseluruhan, maka nilai akurasi untuk strata sebanyak 2,3, dan 4 sama secara keseluruhan yaitu sebesar 0,656 atau 65,6 persen. Hal ini menunjukkan bahwa PSS-SVM dapat memprediksi secara benar pasien DM yang mengalami komplikasi penyakit karena kurang aktivitas olahraga dan pasien DM yang tidak memiliki komplikasi penyakit karena cukup aktivitas olahraga sebesar 65,6 persen.

Tabel 8: Tingkat Akurasi PSS SVM.

Banyak Strata	Strata ke – k	Akurasi strata ke-	Akurasi Keseluruhan
2	1	0,521	0,656
	2	0,792	
3	1	0,563	0,656
	2	0,531	
	3	0,875	
4	1	0,583	0,656
	2	0,458	
	3	0,667	
	4	0,917	

e. *Estimasi Average Treatment Effect (ATE)*

Nilai estimasi ATE ( $\theta$ ) menunjukkan seberapa besar pengaruh variabel aktivitas olahraga (Z) terhadap variabel komplikasi penyakit (Y) ketika pengaruh dari kovariat lain sudah direduksi. Hasil estimasi nilai *Average Treatment Effect (ATE)* dan standar error untuk setiap strata ditunjukkan pada Tabel 9 berikut.

Tabel 9: Estimasi ATE dan Standar Error dengan PSS SVM.

Strata	Strata ke-k	$\hat{\theta}_k$	$\hat{\theta}$	SE ( $\hat{\theta}$ )	p-value
2	1	0,286	0,592	0,129	0,000*
	2	0,307			
3	1	0,142	0,528	0,137	0,000*
	2	0,219			
	3	0,167			
4	1	0,207	0,597	0,103	0,000*
	2	0,108			
	3	0,157			
	4	0,125			

Tabel 9 menunjukkan hasil estimasi efek perlakuan (ATE) untuk PSS SVM dengan  $\hat{\theta}_k$  merupakan estimasi ATE untuk strata ke-k pada strata tertentu,  $\hat{\theta}$  merupakan estimasi keseluruhan ATE dalam strata, dan SE ( $\hat{\theta}$ ) merupakan standar error estimasi ATE. Estimasi efek perlakuan dapat diketahui pada setiap strata, misalnya pada strata sebanyak 2, pada strata ke-1 efek aktivitas olahraga yang cukup terhadap adanya komplikasi penyakit adalah sebesar 0,212. Hal yang sama juga dapat diinterpretasikan untuk strata lainnya.

Berdasarkan Tabel 9 diketahui bahwa masing-masing strata memberikan nilai ATE dan standar error yang hampir sama. Kesimpulan signifikansi dengan taraf signifikansi 0.05 juga memberikan hasil yang sama yakni variabel aktivitas olahraga (Z) merupakan variabel yang berpengaruh terhadap variabel komplikasi penyakit (Y) pada pasien Diabetes Melitus Tipe 2. Jumlah strata 4 memberikan nilai standar error estimasi ATE terkecil yaitu sebesar 0,103 dengan nilai estimasi ATE nya sebesar 0,597.

#### f. Percent Bias Reduction (PBR)

Pada pengujian *balance* kovariat yang telah dilakukan sebelumnya, maka sesuai dengan Tabel 9 diperoleh bahwa strata yang memberikan hasil *balance* untuk semua kovariat berada pada strata sebanyak 2, 3, dan 4. Selanjutnya, untuk menilai apakah metode PSS sudah baik atau tidak, dapat dilihat dari seberapa besar bias yang mampu direduksi oleh metode PSS tersebut. Hasil perhitungan *Percent Bias Reduction* (PBR) ditampilkan pada Tabel 10 berikut.

Tabel 10: Percent Bias Reduction PSS SVM.

Nilai Bias	Banyak Strata		
	2	3	4
<b>Sebelum Strata</b>	0,100	0,100	0,100
<b>Setelah Strata</b>	0,043	0,054	0,014
<b>PBR</b>	57,002	46,371	86,39

Berdasarkan Tabel 10 diketahui bahwa strata yang memberikan reduksi bias terkecil adalah strata sebanyak 3 yaitu sebesar 46,371%, sedangkan reduksi bias terbesar pada PSS SVM yaitu pada strata sebanyak 4 dengan reduksi bias sebesar 86,39 %.

## 4. Simpulan

Penerapan PSS SVM pada kasus komplikasi penyakit pada pasien Diabetes Melitus Tipe 2 diperoleh hasil bahwa variabel yang menjadi *confounding* adalah aktivitas olahraga. Tingkat akurasi PSS SVM adalah sama pada strata sebanyak 2,3, dan 4 sebesar 65,6%. Estimasi efek perlakuan (ATE) memberikan hasil bahwa variabel aktivitas olahraga merupakan variabel yang berpengaruh terhadap komplikasi penyakit (Y) pada pasien Diabetes Melitus Tipe 2. Adapun jumlah strata yang dapat mereduksi bias terbesar yaitu pada strata sebanyak 4 dengan *percent bias reduction* (PBR) sebesar 86,39 % dengan nilai standar error terkecil yaitu sebesar 0,103 dan nilai estimasi ATE sebesar 0,597.

**Ucapan Terima Kasih.** Terima kasih kepada Rumah Sakit Umum Daerah (RSUD) Kabupaten Pasuruan yang telah mengijinkann dilakukanya penelitian dan pengambilan data serta terima kasih kepada pihak pihak lain.

## Daftar Pustaka

Afriza, A. (2015). Pengaruh Senam Diabetes Terhadap Kadar Glukosa Darah pada Penderita Diabetes Mellitus di Puskesmas Lapai Kecamatan Nanggalo Kota Padang. *Jurnal Ilmu Keolahragaan*, 1(1): 1-19.

- Agresti, A. (2003). *Categorical data analysis* (Second). Canada: John Wiley & Sons.
- D'Agostino, R. B. J. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, 17(19): 2265–2281. [https://doi.org/10.1002/\(sici\)1097-0258\(19981015\)17:19<2265::aid-sim918>3.0.co;2-b](https://doi.org/10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b)
- Gungor, N., Hannon, T., Libman, I., Bacha, F., & Arslanian, S. (2005). Type 2 diabetes mellitus in youth: the complete picture to date. *Pediatric Clinics of North America*, 52(6): 1579–1609. <https://doi.org/10.1016/j.pcl.2005.07.009>
- Härdle, W. K., Prastyo, D. D., & Hafner, C. M. (2014). Support vector machines with evolutionary model selection for default prediction. In *The oxford handbook of applied nonparametric and semiparametric econometrics and statistics* (pp. 346–373). <https://doi.org/10.1093/oxfordhb/9780199857944.013.011>
- [KEMENKES RI] Kementerian Kesehatan RI. (2013). *Infodatin Pusat Data dan Informasi Kemenkes RI: Situasi dan Analisis Diabetes*. Jakarta (ID): Kementerian Kesehatan RI.
- Li, M. (2013). Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods*, 16(2): 188–226.
- McCandless, L. C., Gustafson, P., & Austin, P. C. (2009). Bayesian propensity score analysis for observational data. *Statistics in Medicine*, 28(1): 94–112.
- Merdeka, M. B. A. (2017). *Propensity Score Stratification Menggunakan Regresi Logistik Pada Kasus Infeksi Saluran Pernapasan Akut Terhadap Balita Di Puskesmas Pasuruan Jawa Timur*. Surabaya (ID): Institut Teknologi Sepuluh November.
- Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4): 603–614.
- Misnadiarly. (2006). *Diabetes mellitus gangren, ulcer, infeksi: mengenali gejala, menanggulangi, mencegah komplikasi*. Jakarta (ID): Pustaka Populer Obor.
- Pan, W., & Bai, H. (2015). *Propensity score analysis: Fundamentals and developments*. Guilford Publications.
- Provost, F. (1998). Glossary of Terms Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. *Machine Learning*, 30: 271–274.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1): 41–55.
- Sharoh, S. M., & Salmiyati, S. (2017). *Pengaruh Senam Diabetes terhadap Kadar Gula Darah pada Penderita Diabetes Mellitus Tipe II di Wilayah Kerja Puskesmas Gamping 1 Sleman Yogyakarta*. Universitas.
- Sinaga, J., & Hondro, E. (2011). Pengaruh senam diabetes mellitus terhadap kadar glukosa darah pada penderita diabetes mellitus tipe 2 di wilayah kerja puskesmas darusalam medan. *Jurnal Mutiara Ners*, 1(7): 1-7.

- Sustrani, L., Alam, S., & Hadibroto, I. (2006). *Diabetes: Informasi Lengkap untuk Penderita dan Keluarganya*. Jakarta (ID): Gramedia Pustaka.
- Tu, W., & Zhou, X.-H. (2002). A Bootstrap Confidence Interval Procedure for the Treatment Effect Using Propensity Score Subclassification. *Health Services and Outcomes Research Methodology*, 3(2): 135–147. <https://doi.org/10.1023/A:1024212107921>
- Westreich, D., Lessler, J., & Funk, M. J. (2010). Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology*, 63(8): 826–833. <https://doi.org/10.1016/j.jclinepi.2009.11.020>
- Winkelmayer, W. C., & Kurth, T. (2004). Propensity scores: help or hype? *Nephrology Dialysis Transplantation*, 19(7): 1671–1673. <https://doi.org/10.1093/ndt/gfh104>
- [WHO] World Health Organization. (2015). Jakarta (ID): *WHO Statistical Profile*.
- Yu, H., Hong, S., Yang, X., Ni, J., Dan, Y., & Qin, B. (2013). Recognition of Multiple Imbalanced Cancer Types Based on DNA Microarray Data Using Ensemble Classifiers. *BioMed Research International*, 2013: 1-13. <https://doi.org/10.1155/2013/239628>
- Zhao, C. Y., Zhang, H. X., Zhang, X. Y., Liu, M. C., Hu, Z. D., & Fan, B. T. (2006). Application of support vector machine (SVM) for prediction toxic activity of different data sets. *Toxicology*, 217(2): 105–119. <https://doi.org/10.1016/j.tox.2005.08.019>