

Received: 25 April 2019

Revised: 14 Mei 2019

Accepted: 22 Mei 2019

Published: 30 Juni 2019

FAKTOR-FAKTOR YANG MEMENGARUHI PERNIKAHAN DINI PADA WANITA USIA 20-24 DI INDONESIA TAHUN 2017: PENERAPAN METODE REGRESI LOGISTIK BINER DENGAN PENYESUAIAN *RESAMPLING* DATA *IMBALANCE*

Ersa Budi Sutanto^{1, a)}, Ghytsa Alif Jabir^{1, b)}, Nadhifan Humam Fitriah^{1, c)}, Ni Luh Putu Yayang
Septia Ningsih^{1, d)}, Siti Andhasah^{1, e)}, Rani Nooraeni^{1, f)}

¹ *Politeknik Statistika STIS, Jl. Otto Iskandardinata No.64C Jakarta 13330.*

Email: ^{a)}211709664@stis.ac.id, ^{b)}211709715@stis.ac.id, ^{c)}211709877@stis.ac.id, ^{d)}211709892@stis.ac.id,
^{e)}211710015@stis.ac.id, ^{f)}raninoor@stis.ac.id

Abstract

Child marriage is a marriage that occurs in children under the age of 18 years. In general, the prevalence of child marriage in Indonesia is still quite high because 23 of 34 provinces in Indonesia have a prevalence of marriage at the age of children above the national average. The category used in this case shows the imbalance, so it is necessary to adjust in analyzing the data. Problems that are often encountered data access cases of child marriage are not recorded or recorded on official documents. To overcome this, this study uses the recognition of women aged 20-24 years at the time of the survey which stated that they had married before the age of 18 years. The variables used in this study were poverty status, classification of residential areas, early marital status of household heads (KRT), KRT education, type of household work, number of household members, and internet usage status of women aged 20-24 years. Before being analyzed, the imbalance data was adjusted by resampling techniques which included: over sampling, under sampling, and hybrid then analyzed using binary logistic regression methods. It will also look for factors that influence child marriage in women aged 20-24 years. Based on the results of processing with the R program, the most appropriate resampling technique for this case is oversampling. With this resampling method, all variables significantly influence child marriage in women aged 20-24.

Keywords: child marriage, imbalance, resampling, biner logistic

Abstrak

Pernikahan dini adalah perkawinan yang terjadi pada anak dibawah usia 18 tahun. Secara umum, angka prevalensi pernikahan dini di Indonesia masih cukup tinggi karena 23 dari 34 provinsi di Indonesia memiliki angka prevalensi pernikahan pada usia dini di atas rata-rata nasional. Kategori yang digunakan pada kasus ini menunjukkan keadaan *imbalance* sehingga diperlukan adanya penyesuaian dalam

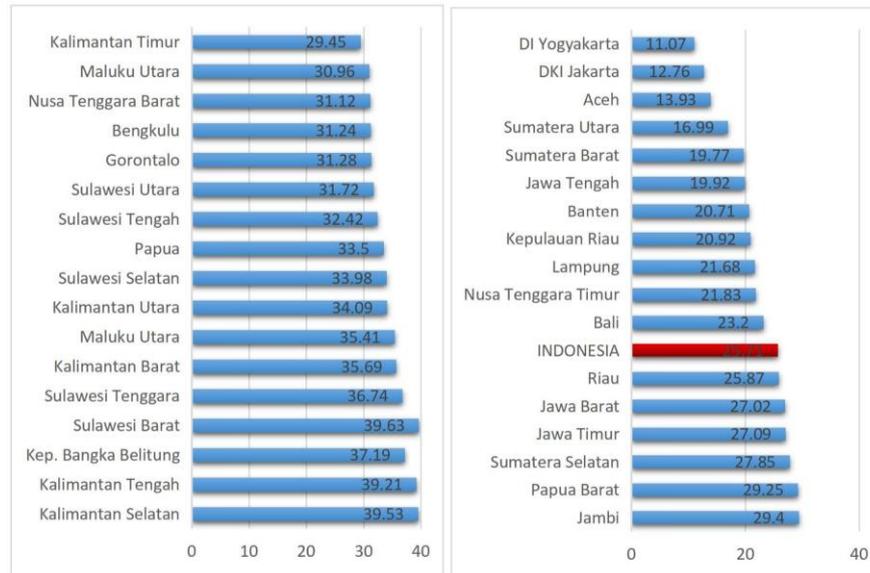
menganalisis data. Permasalahan yang sering dijumpai akses data kasus pernikahan dini tidak tercatat atau terekam pada dokumen resmi. Dalam upaya untuk mengatasi hal tersebut, penelitian ini menggunakan pengakuan dari perempuan berusia 20-24 tahun pada saat dilakukan survei yang menyatakan bahwa mereka pernah kawin sebelum usia 18 tahun. Variabel yang digunakan dalam penelitian ini yaitu status kemiskinan, klasifikasi wilayah tempat tinggal, status pernikahan dini kepala rumah tangga (KRT), pendidikan KRT, jenis pekerjaan KRT, jumlah anggota rumah tangga, dan status penggunaan internet perempuan berusia 20-24 tahun. Sebelum dianalisis, data *imbalance* tersebut dilakukan penyesuaian dengan teknik *resampling* yang meliputi: *oversampling*, *undersampling*, dan *hybrid* kemudian dianalisis dengan metode *regresi logistik biner*. Selain itu juga akan dicari faktor-faktor yang memengaruhi pernikahan dini pada perempuan berusia 20-24 tahun. Berdasarkan hasil pengolahan dengan program R, teknik *resampling* yang paling tepat untuk kasus ini adalah *oversampling*. Dengan metode *resampling* tersebut, seluruh variabel signifikan berpengaruh terhadap pernikahan dini pada perempuan usia 20-24.

Kata-kata kunci: pernikahan dini, *imbalance*, *resampling*, logistik biner

PENDAHULUAN

Salah satu target *Sustainable Development Goals* (SDG's) dalam tujuan kelima adalah meraih kesetaraan gender dan pemberdayaan seluruh perempuan dan anak-anak perempuan. Untuk meraih tujuan tersebut, hal yang dapat dilakukan adalah dengan mengurangi praktik-praktik berbahaya pada anak-anak, termasuk diantaranya pernikahan pada usia dini. Hal ini karena pernikahan dimaksudkan untuk membina hubungan yang langgeng antara kedua pasangan, sehingga dalam menjalaninya dibutuhkan kedewasaan dan tanggung jawab antara kedua mempelai baik secara fisik maupun batin (mental).

Dalam istilah internasional pernikahan dini dikenal dengan *child marriage* atau *early marriage*, adalah perkawinan yang terjadi pada anak dibawah usia 18 tahun. Pembatasan dalam angka 18 ini sesuai dengan batas usia perlindungan anak yang ditetapkan dalam konvensi hak-hak anak internasional (*Convention on The Rights of The Child*) pada tahun 1989 (Justice for Iran, 2013). Sedangkan usia ideal menikah pada perempuan yaitu berada pada 21 tahun atau lebih dan pada laki-laki pada usia 25 tahun atau lebih (Bkkbn, 2010). Pada usia tersebut organ reproduksi khususnya bagi perempuan sudah berkembang secara baik dan siap untuk melahirkan. Begitu pula pada laki-laki, pada usia tersebut seorang laki-laki cenderung lebih siap untuk menopang kehidupan keluarganya. Namun dalam prakteknya, pernikahan pada usia dini masih banyak dijumpai di negara berkembang, salah satunya indonesia. Hal ini dapat disebabkan karena sebagian besar warga negara di negara berkembang belum menyadari bahaya kehamilan atau melahirkan pada ibu yang berumur kurang dari 20 tahun.



Sumber: Badan Pusat Statistik

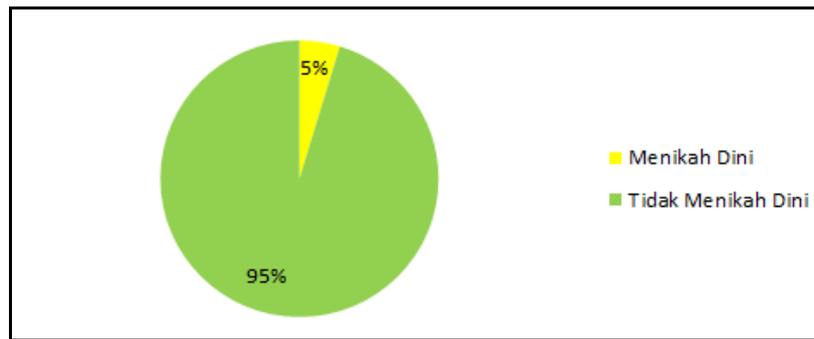
Gambar 1. Persentase Perempuan 20-24 Tahun dengan Umur Kawin Pertama dibawah 18 Tahun Menurut Provinsi Tahun 2017

Badan Pusat Statistik (BPS) menyebutkan, di dunia terdapat lebih dari 700 juta perempuan menikah sebelum usianya mencapai 18 tahun, bahkan sekitar 250 juta diantaranya menikah dibawah usia 15 tahun. Sedangkan di Indonesia, BPS dan UNICEF mencatat terdapat sekitar 340 ribu anak perempuan dibawah 18 tahun menikah setiap tahunnya. Peningkatan terjadi khususnya pada perempuan usia 15-18 tahun.

Pada tahun 2017 sendiri, secara rata-rata angka prevalensi pernikahan wanita dibawah usia 18 tahun di Indonesia berada di angka 25,71 persen. Dengan kata lain, terdapat satu dari empat wanita di Indonesia telah menikah dibawah usia 18 tahun. Prevalensi tertinggi berada di provinsi Kalimantan Selatan dengan angka 39,53 persen sedangkan terendah berada di provinsi Yogyakarta dengan angka prevalensi sebesar 11,07 persen. Secara umum kawasan Indonesia bagian timur memiliki angka prevalensi lebih tinggi dibandingkan dengan kawasan Indonesia bagian barat. Jika diamati lebih lanjut, 23 dari 34 provinsi di Indonesia memiliki angka prevalensi pernikahan pada usia dini diatas rata-rata nasional. Angka ini tergolong cukup tinggi dan mengkhawatirkan karena peningkatan angka pernikahan usia dini dapat berdampak kepada nasib wanita di Indonesia.

Permasalahan yang sering dijumpai di lapangan adalah akses data mengenai status pernikahan atau usia pernikahan pada wanita usia dini yang tidak mudah diperoleh karena terlalu banyak kasus pernikahan dini tidak tercatat atau terekam pada dokumen resmi. Maka, dalam upaya untuk menghitung prevalensi pernikahan dini secara tepat, penelitian ini menggunakan pengakuan dari perempuan berusia 20-24 tahun pada saat dilakukan survei yang menyatakan bahwa mereka pernah kawin sebelum usia 18 tahun.

Pada penelitian ini, populasi yang menjadi objek penelitian merupakan seluruh wanita usia 20-24 tahun berstatus anak dalam rumah tangga. Populasi ini mencakup wanita yang berstatus kawin maupun belum kawin, sehingga menyebabkan kategori wanita yang berstatus nikah dini dengan tidak nikah dini menjadi tidak seimbang.



Gambar 2. Persentase Pernikahan Dini Perempuan 20-24 Tahun

Berdasarkan grafik di atas, persentase wanita berusia 20-24 tahun yang melakukan pernikahan dini hanya sebesar 5%, sedangkan persentase yang tidak melakukan pernikahan dini sebesar 95%. meliputi wanita yang melakukan pernikahan pada usia diatas 18 tahun dan wanita yang belum menikah sampai mencapai usia 20-24 tahun. Hal tersebut menunjukkan adanya ketimpangan antara wanita yang melakukan pernikahan dini dan tidak melakukan pernikahan dini sehingga menimbulkan imbalanced data. Kasus *imbalanced* terjadi ketika salah satu kategori memiliki jumlah proporsi yang lebih besar dibandingkan kategori lainnya. Apabila dipaksakan menggunakan regresi logistik klasik tanpa *treatment*, maka model penaksir parameter yang terbentuk akan bias dan menghasilkan klasifikasi yang tidak tepat (*overestimate* pada kategori yang memiliki proporsi yang besar dan klasifikasi yang *underestimate* pada kategori yang memiliki proporsi yang kecil).

Oleh karena itu, sebelum masuk tahap pengolahan, data akan diseimbangkan dengan menggunakan beberapa teknik *resampling* yaitu *over sampling*, *under sampling*, dan *hybrid*. Kemudian hasil pengolahan teknik *resampling* akan dibandingkan untuk mendapat teknik *resampling* yang terbaik. Selain itu juga akan dicari faktor-faktor yang memengaruhi pernikahan dini pada perempuan berusia 20-24 tahun dengan menggunakan metode analisis *regresi logistik biner*.

METODE PENELITIAN

Ruang Lingkup Penelitian

Pada penelitian ini, pengumpulan data dilakukan dengan menggunakan data sekunder yang bersumber dari *raw data* Survei Sosial Ekonomi Nasional (Susenas) tahun 2017 yang dilakukan oleh Badan Pusat Statistik. Susenas menyediakan data berbagai aspek sosial ekonomi dan pemenuhan kebutuhan hidup seperti sandang, pangan, papan, pendidikan, kesehatan, keamanan, dan kesempatan kerja.

Metode Resampling

Pada penelitian ini akan dilakukan perbandingan antara tiga metode *resampling* yang biasa digunakan dalam mengatasi *imbalanced data* atau *imbalanced class*. Teknik *resampling* adalah proses menyeimbangkan distribusi data kembali untuk mengurangi efek distribusi kelas yang tidak seimbang dalam proses data training, dimana data asli yang tidak seimbang tersebut akan diseimbangkan berdasarkan serangkaian metode algoritma sampling yang menyesuaikan dengan jumlah kelas yang berbeda.

Metode *resampling* yang pertama adalah metode *oversampling*, dimana cara kerja metode ini adalah dengan menduplikasi data dari kelas minoritas yang dipilih secara acak sehingga jumlah kelas pada data kelas minoritas akan meningkat. Salah satu teknik untuk menangani kasus *imbalanced data* atau *imbalanced class* yaitu *oversampling*. Metode *resampling* yang kedua adalah metode *undersampling*, dimana cara kerja metode ini adalah mengeliminasi data sampel yang terdapat pada kelas mayoritas sehingga distribusi data sampel menjadi lebih seimbang. Metode *resampling* yang

ketiga adalah metode *Hybrid* yaitu metode yang menggabungkan antara metode *oversampling* dan *undersampling*. Pemilihan metode resampling yang tepat pada penelitian ini didasarkan pada ukuran performa yaitu *sensitivity*, *precision*, dan *balanced accuracy*.

Salah satu cara untuk membandingkan performa dari model yang telah dibuat adalah dengan menggunakan *confusion matrix*. *Confusion matrix* merupakan tabel yang merepresentasikan nilai sebenarnya dari data asli dengan nilai prediksi dari model yang telah dibuat. Berikut contoh *confusion matrix* pada kasus klasifikasi biner.

TABEL 1. Confusion Matrix

		Actual Class	
		0	1
Predicted Class	0	TN	FN
	1	FP	TP

True Positive (TP) dan *True Negative (TN)* merupakan jumlah observasi yang dapat diklasifikasikan dengan benar oleh model. Sedangkan *False Negative (FN)* dan *False Positive* merupakan jumlah observasi yang tidak diklasifikasikan dengan benar oleh model. Dari *confusion matrix* ini dapat dihitung berbagai ukuran performa seperti *Sensitivity/Recall*, *Precision*, dan *Balanced Accuration*.

Sensitivity/Recall merupakan proporsi antara prediksi observasi positif atau kategori “1” terhadap banyaknya observasi positif sebenarnya atau kategori “1”. Dengan kata lain, *Sensitivity/Recall* merupakan proporsi kategori positif yang dapat diklasifikasikan dengan benar oleh model. *Sensitivity/Recall* dapat dirumuskan sebagai berikut.

$$sensitivity = \frac{TP}{TP + FN} \tag{1}$$

Precision merupakan proporsi antara prediksi observasi positif atau kategori “1” yang benar terhadap banyaknya prediksi observasi positif. Dengan kata lain, *precision* merupakan proporsi prediksi observasi positif yang dapat diklasifikasikan dengan benar oleh model. *Precision* dapat dirumuskan sebagai berikut.

$$precision = \frac{TP}{TP + FP} \tag{2}$$

Balanced Accuracy merupakan proporsi tertimbang antara observasi yang dapat diprediksi dengan benar oleh model pada setiap kelas. *Balanced Accuration* dapat dirumuskan:

$$balanced_accuracy = \frac{TP}{TP + FP} + \frac{TN}{TN + FP} \tag{3}$$

Regresi Logistik Biner

Menurut Hosmer dan Lemeshow (2000) dalam *Applied Logistic Regression*, metode regresi logistik biner merupakan metode yang digunakan untuk menggambarkan hubungan antara variabel respons (*dependent variable*) dengan satu atau lebih variabel penjelas (*independent variables*), dimana variabel tak bebasnya diasumsikan berkategori diskrit dengan dua kemungkinan nilai yaitu “sukses” dan “gagal”.

Bentuk umum dari Model Regresi Logistik Biner adalah sebagai berikut.

$$Li = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (4)$$

Uji Simultan

Uji *Omnibus test* dapat digunakan untuk mengetahui pengaruh variabel penjelas terhadap variabel respons secara simultan atau bersama-sama. Hipotesis yang digunakan untuk uji ini adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Minimal terdapat satu } \beta_j \neq 0$$

Pengambilan keputusan didasarkan pada nilai *p-value*. Apabila *p-value* < α , maka tolak tolak H_0 . Artinya, dapat dinyatakan bahwa dengan tingkat signifikansi sebesar α , minimal terdapat satu variabel independen yang memengaruhi variabel dependen.

Uji Parsial

Uji parsial dilakukan guna mengetahui pengaruh variabel penjelas terhadap variabel respons dalam model secara parsial. Hipotesis yang digunakan dalam uji ini adalah sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Statistik uji *Wald* digunakan dalam pengujian signifikansi variabel, dengan rumus:

$$W_j = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2 \sim \chi^2_{(p)} \quad (5)$$

Statistik uji *Wald* (W_j) mengikuti distribusi *Chi-Squares* (χ^2) dengan derajat bebas satu, sehingga pengambilan keputusan didasarkan pada Nilai $W_j > \chi^2_{(\alpha,1)}$ atau *p-value* < α , maka tolak H_0 . Artinya, dengan tingkat signifikansi sebesar α , terdapat pengaruh signifikan antara variabel independen ke- j terhadap variabel dependen.

Uji Kesesuaian Model (*Goodness of Fit Test*)

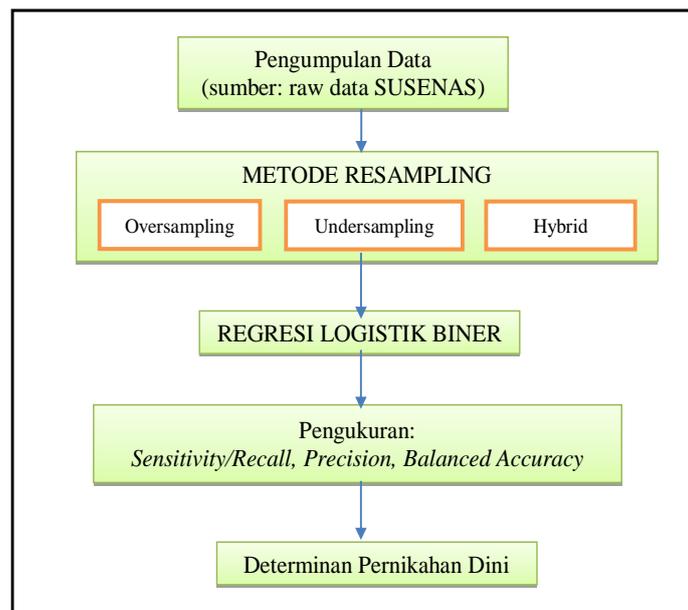
Uji kesesuaian model atau *goodness of fit test* dilakukan untuk memeriksa apakah ada perbedaan antara hasil yang diperoleh dari model dengan hasil yang diharapkan dari data observasi. Pengujian kesesuaian model dalam regresi logistik dilakukan dengan menggunakan *The Hosmer and Lemeshow Test*. Hipotesis uji yang digunakan adalah sebagai berikut.

$$H_0 : \text{Model fit}$$

$$H_1 : \text{Model tidak fit}$$

Pengambilan keputusan didasarkan pada nilai *p-value*. Apabila *p-value* < α , maka tolak tolak H_0 . Namun dalam uji kesesuaian model, hasil yang diharapkan adalah gagal tolak H_0 , agar dapat diambil kesimpulan bahwa model yang terbentuk adalah fit atau model yang terbentuk sudah cocok.

Secara umum, metodologi yang digunakan dalam penelitian akan ditampilkan dalam diagram berikut.



Gambar 3. Alur Metodologi Penelitian

Kajian Pustaka

Variabel *dependen* dalam penelitian ini adalah perempuan usia 20-24 tahun yang melakukan pernikahan dini. Adapun variabel *independen* yang digunakan yaitu status kemiskinan, klasifikasi wilayah tempat tinggal, status pernikahan dini kepala rumah tangga (KRT), pendidikan KRT, jenis pekerjaan KRT, jumlah anggota rumah tangga, dan status penggunaan internet perempuan berusia 20-24 tahun.

1. Pernikahan Dini

Pernikahan dini adalah pernikahan yang terjadi pada seseorang yang belum memiliki cukup usia standar, dalam hal ini konvensi Hak-hak Anak International menyatakan bahwa usia standar pernikahan dini yaitu dibawah usia 18 tahun.

2. Status Kemiskinan

Faktor yang dianggap paling berpengaruh terhadap pernikahan dini yaitu faktor ekonomi yang didekati dengan status kemiskinan. Untuk menunjukkan status kemiskinan seseorang, BPS menggunakan konsep penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan. Sebagian besar orang berpendapat bahwa ketika perempuan sudah menikah perekonomian keluarga menjadi lebih baik karena mendapat nafkah dan jaminan ekonomi dari suami. Selain itu, menurut Stang (2011), responden yang berpendapatan rendah lebih banyak melakukan pernikahan dini dikarenakan masih ada beberapa orang yang beralasan menikah yang didukung oleh keadaan ekonomi yang rendah, sehingga orang tua menyuruh anaknya menikah saja untuk mengurangi beban rumah tangga dan mandiri.

3. Klasifikasi Wilayah Tempat Tinggal

Dalam penelitian ini klasifikasi wilayah tempat tinggal terbagi menjadi dua, yaitu perkotaan dan perdesaan. Kedua daerah tersebut cenderung memiliki karakteristik yang berbeda. Baik dari segi fasilitas maupun infrastruktur yang dapat mempengaruhi keputusan seseorang untuk menikah dini. Menurut Mariyatul Qibtiyah (2014), responden yang tinggal di desa cenderung menikah di usia muda daripada di kota, yaitu sebesar 52,9%. Hal ini didukung oleh penelitian Soekarno (2012) dalam Mariyatul Qibtiyah (2011) yang menyatakan bahwa rata-rata umur kawin pertama lebih rendah di wilayah perdesaan dibandingkan perdesaan.

4. Jumlah Anggota Rumah Tangga

Adapun jika dilihat berdasarkan jumlah anggota rumah tangga, menurut Subidia,dkk (2015), terdapat hubungan negatif antara jumlah anggota rumah tangga dengan usia kawin pertama. Artinya, semakin banyak anggota rumah tangga, maka usia kawin pertamanya semakin rendah. Hal ini sejalan dengan dugaan penelitian ini yang dapat diasumsikan bahwa semakin banyak anggota rumah tangga, maka kecenderungan untuk menikah dini semakin tinggi.

5. Status Pernikahan Dini KRT

Faktor lain yang mempengaruhi pernikahan dini adalah faktor budaya. Di beberapa wilayah Indonesia, pernikahan dini dianggap hal yang tabu bahkan sebuah keharusan. Menurut Agus Mahfudin, dkk. (2016) dalam penelitian kualitatif yang dilakukan di daerah Kabupaten Sumenep Jawa Timur, terdapat responden yang menyatakan bahwa ia menikah muda dengan alasan adat dalam keluarga yang sudah turun temurun dan sudah tidak bisa ditentang lagi. Hal tersebut mengakibatkan sang anak mau tidak mau harus mengikuti riwayat pernikahan orang tuanya. Sehingga kecenderungan seseorang yang menikah dini juga dipengaruhi status menikah dini sang orang tua.

6. Pendidikan KRT

Peran orang tua terhadap kelangsungan pernikahan dini tidak dapat terlepas dari tingkat pengetahuan orang tua yang dihubungkan dengan pendidikan orang tua atau kepala rumah tangga. Apalagi sebagian besar anak pasti dididik dan dibesarkan oleh orang tua kandungnya. Menurut Nandang,dkk (2009) dalam Irne (2015) menyatakan bahwa terdapat hubungan antara pendidikan orang tua pada wanita dewasa muda. Remaja yang memiliki latar belakang orang tua berpendidikan rendah maka memiliki resiko sebesar 7,667 kali lebih besar untuk menikah dini daripada remaja yang memiliki latarbelakang orang tua berpendidikan tinggi.

7. Status Pekerjaan KRT

Menurut Desyianti (2015) dalam Siti Salamah (2016), faktor yang mempengaruhi kejadian pernikahan dini bukan dari sudut pekerjaan responden melainkan pekerjaan orang tua. Selain itu, menurut Yunita (2014), dengan pekerjaan orang tua maka akan mencerminkan status sosial ekonomi dari keluarga remaja tersebut. Ketika seorang perempuan berusia dibawah 18 tahun, asumsi awal peneliti adalah wanita tersebut belum bekerja karena belum memasuki batas usia minimum pekerja perempuan dan masih menjadi tanggungan orang tua. Sesuai dengan UU No.13 tahun 2003 yang menyatakan bahwa seseorang dilarang mempekerjakan perempuan yang berumur kurang dari 18 (delapan belas) tahun. Sehingga proksi variabel yang lebih cocok untuk digunakan adalah pekerjaan orang tua.

8. Status Penggunaan Internet

Pada sisi lain, era kemajuan berkembang dengan pesat, salah satunya didukung dengan kemudahan akses internet. Menurut Sarwono (2012), Internet adalah suatu kumpulan jaringan yang memiliki skala global. Menurutnya tidak ada orang yang mampu bertanggung jawab untuk menggunakan internet itu. Pengaruh teknologi seperti internet memberikan kemudahan mengakses internet dengan cepat. Selain itu, akses internet yang tak terbatas telah membawa pengaruh besar di segala lini yang salah satunya adalah pernikahan dini. Melalui internet, seseorang yang memiliki usia dibawah umur dapat dengan mudah menggunakan internet untuk mengakses film porno dari *website* ilegal. Anak usia dibawah umur yang memiliki kemudahan akses internet beresiko lebih besar untuk menikah dini dibandingkan dengan anak dibawah umur yang jarang menggunakan internet.

HASIL DAN PEMBAHASAN

Perbandingan dan Pemilihan Metode *Resampling*

Sebelum dilakukan pengolahan, data dibagi menjadi dua jenis, yakni *train data* dan *test data* dengan perbandingan masing-masing 70% dan 30% proporsional tiap jenis data. Metode *resampling*

hanya diterapkan pada *train data*, Pada *train data* diterapkan tiga metode *resampling* yang berbeda, yakni: *undersampling*, *oversampling*, dan *hybrid*. Tiap data yang telah diterapkan metode *resampling* digunakan untuk membentuk model regresi logistik biner. Model kemudian akan dievaluasi menggunakan data *test data* untuk diketahui performa masing-masing metode *resampling* dengan kriteria antara lain adalah *sensitivity*, *specificity*, *precision*, dan ROC. Berikut adalah tabel performa masing-masing metode *resampling*.

TABEL 2. Akurasi Model

Performa Pengukuran	Tanpa Resampling	Under Sampling	Over Sampling	Hybrid
<i>Sensitivity</i>	0,000	0,722	0,731	0,715
<i>Specificity</i>	0,000	0,107	0,108	0,109
<i>Balanced Accuracy</i>	0.500	0,713	0,717	0,713

Berdasarkan tabel perbandingan *sensitivity*, *precision*, dan *balanced accuracy*, dapat dilihat bahwasanya pada metode *over sampling* memiliki nilai *sensitivity* paling besar, yakni sebesar 0.73463. Dalam penelitian ini yang menjadi fokus utama adalah memprediksi data minoritas yaitu wanita yang menikah dini secara tepat. Nilai *sensitivity* adalah banyak wanita menikah dini yang diklasifikasikan secara tepat oleh model dibagi dengan banyak wanita menikah dini. Selanjutnya berdasarkan nilai *precision*, metode *over sampling* berada para urutan kedua setelah metode *hybrid*. Nilai *precision* merupakan banyak wanita menikah dini yang diklasifikasikan oleh model secara tepat dibagi banyak wanita menikah dini yang diprediksi menikah dini. Selain itu, digunakan performa *balanced accuracy* atau akurasi tertimbang, didapat nilai dari metode *over sampling* paling tinggi di antara metode *resampling* lainnya. Berdasarkan ketiga pengukuran performa, diputuskan metode *resampling* terbaik untuk kasus ini adalah metode *over sampling*.

Variabel-variabel yang Memengaruhi Kejadian Menikah Dini pada Perempuan Berusia 20 – 24 Tahun

Untuk mengetahui apakah variabel independen yang meliputi status kemiskinan, klasifikasi wilayah tempat tinggal, status nikah dini kepala rumah tangga (KRT), pendidikan KRT, jenis pekerjaan KRT, jumlah anggota rumah tangga, dan status penggunaan internet individu tersebut secara bersama-sama berpengaruh signifikan terhadap variable dependen (status menikah dini perempuan usia 20-24 tahun), maka dilakukan pengujian secara simultan.

Berdasarkan uji Omnibus menggunakan *software R*, didapatkan *p-value* sebesar 0,000, Artinya, dengan taraf signifikansi 5% terdapat cukup bukti bahwa status kemiskinan, klasifikasi wilayah tempat tinggal, status nikah dini kepala rumah tangga (KRT), pendidikan KRT, jenis pekerjaan KRT, jumlah anggota rumah tangga, dan status penggunaan internet individu tersebut secara bersama-sama berpengaruh signifikan terhadap variabel dependen (status menikah dini perempuan usia 20-24 tahun).

Setelah dilakukan uji parameter simultan, maka selanjutnya dapat dilakukan uji parameter parsial untuk mengetahui variable bebas mana yang berpengaruh sevara signifikan terhadap terjadinya pernikahan dini pada perempuan usia 20-24 tahun. Berikut adalah tabel parameter hasil uji parsial dengan menggunakan *Wald Test*. Berikut adalah hasil uji parsial dengan menggunakan *Wald Test*.

TABEL 3. Hasil Uji Parsial (*Wald Test*)

Variabel	Kategori	$\hat{\beta}$	<i>p-value</i>	$Exp(\hat{\beta})$
<i>Intercept</i>		-1,795	0,000	0,166
Status Kemiskinan (D1)	1 = miskin 0 = tidak miskin*	0,098	0,010	1,104
Klasifikasi Wilayah (D2)	1 = perdesaan 0 = perkotaan*	0,187	0,000	1,206
Jumlah ART (D3)		0,278	0,000	1,320
Pernikahan Dini KRT (D4)	1 = menikah dini 0 = tidak menikah dini*	0,927	0,000	2,527
Pendidikan KRT (D5)	1 = ≤SMP 0 = >SMP*	0,851	0,000	2,342

Status Pekerjaan KRT (D6)	1 = tidak bekerja 0 = bekerja*	0.206	0,000	1,229
Penggunaan Internet (D7)	1 = menggunakan internet 0 = tidak menggunakan internet*	-1.254	0,000	0,285

Keterangan = tanda * merupakan kategori *reference*

Berdasarkan hasil pengujian parameter secara parsial, diketahui bahwa semua variabel yaitu status kemiskinan, klasifikasi wilayah, jumlah ART, status pernikahan dini KRT, pendidikan KRT, status pekerjaan KRT, dan status penggunaan internet berpengaruh secara signifikan terhadap terjadinya pernikahan dini pada perempuan usia 20-24 tahun. Persamaan model regresi logistik yang terbentuk adalah,

$$Li = \ln\left(\frac{p_i}{1 - p_i}\right) = -1.795 + 0.098D_{11}^* + 0.187D_{21}^* + 0.278D_3^* + 0.927D_{41}^* + 0.851D_{51}^* + 0.206D_{61}^* - 0.098D_{71}^*$$

Berdasarkan persamaan di atas dapat diketahui bahwa perempuan yang berusia 20-24 tahun yang berstatus tidak miskin, tinggal di wilayah perkotaan, KRT tidak menikah dini, pendidikan tertinggi yang pernah ditamatkan oleh KRT adalah lebih dari SMP dan status pekerjaannya adalah bekerja yaitu 0,166 kali lebih kecil untuk melakukan pernikahan dini. Berdasarkan status kemiskinannya, perempuan berusia 20-24 tahun yang berstatus miskin memiliki kecenderungan 1.104 kali lebih besar untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang berstatus tidak miskin. Berdasarkan klasifikasi wilayahnya, perempuan berusia 20-24 tahun yang tinggal di wilayah pedesaan memiliki kecenderungan 1,206 kali lebih besar untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang tinggal di wilayah perkotaan. Berdasarkan jumlah anggota rumah tangganya, perempuan berusia 20-24 tahun akan memiliki kecenderungan 1,32 kali lebih besar untuk melakukan pernikahan dini pada setiap kenaikan satu anggota rumah tangga perempuan tersebut. Berdasarkan status pernikahan dini KRT, perempuan berusia 20-24 tahun yang yang KRT-nya melakukan pernikahan dini memiliki kecenderungan 2,527 kali lebih besar untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang KRT-nya tidak melakukan pernikahan dini. Berdasarkan pendidikan KRT, perempuan berusia 20-24 tahun yang memiliki KRT dengan tingkat pendidikan maksimal SMP memiliki kecenderungan 2,342 kali lebih besar untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang memiliki KRT dengan tingkat pendidikan lebih dari SMP. Berdasarkan pekerjaan KRT, perempuan berusia 20-24 tahun yang KRT-nya tidak bekerja memiliki kecenderungan 1,229 kali lebih besar untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang KRT-nya bekerja. Berdasarkan status penggunaan internet, perempuan berusia 20-24 tahun yang menggunakan internet memiliki kecenderungan 0,285 kali lebih kecil untuk melakukan pernikahan dini dibandingkan dengan perempuan berusia 20-24 tahun yang tidak menggunakan internet.

Uji Kesesuaian Model

Uji Kesesuaian Model diperlukan untuk mengetahui apakah suatu model sudah sesuai atau tidak dalam menjelaskan kejadian pernikahan dini pada perempuan berusia 20-24 tahun. Uji ini menggunakan *Hosmer dan Lemeshow Test*. Hasil pengujiannya menunjukkan *p-value* 0.9998 sehingga keputusan yang diambil adalah gagal tolak H_0 . Keputusan tersebut bermakna model yang digunakan sesuai (fit) dalam menjelaskan terjadinya pernikahan dini pada perempuan usia 20-24 tahun.

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa pada kasus pernikahan dini dengan menggunakan data SUSENAS 2017, metode *resampling* yang paling tepat untuk mengatasi data *imbalanced* adalah *oversampling*. Hal ini didasarkan pada nilai *sensitivity*, *precision*, dan *balanced*

accuracy metode *oversampling* yang paling tinggi jika dibandingkan dengan metode *undersampling* dan *hybrid*.

Kemudian, hasil dari analisis regresi logistik biner dapat ditunjukkan bahwa variabel-variabel yang signifikan dalam mempengaruhi perempuan berusia 20-24 tahun untuk melakukan pernikahan dini adalah status kemiskinan, klasifikasi wilayah, jumlah ART, status pernikahan dini KRT, pendidikan KRT, status pekerjaan KRT, dan status penggunaan internet.

UCAPAN TERIMA KASIH

Terima kasih kepada Ibu Dosen pembimbing kami yang telah memberi saran dan masukan serta terimakasih kepada teman-teman yang telah menyumbangkan tenaga dan pikiran sehingga penelitian ini dapat selesai dengan lancar.

REFERENSI

- Baizal, Z.A., Bijaksana, M.A. and Sastrawan, A.S. (2009) 'Analisis pengaruh metode oversampling dalam churn prediction untuk perusahaan telekomunikasi', In *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*.
- Desiyanti, I.W. (2015) 'Faktor-faktor yang berhubungan terhadap pernikahan dini pada pasangan usia subur di Kecamatan Mapanget Kota Manado', *Jikmu*, 5(3), pp.270-280.
- Heranova, O. (2019) 'Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring', *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 3(3), pp.443-450.
- Ketut, S.I., IGA, M.D. and Nyoman, D.R.I. (2015) 'FAKTOR-FAKTOR YANG MEMPENGARUHI MENURUNNYA USIA KAWIN PERTAMA DI PROVINSI BALI', *PIRAMIDA*, 11(2), pp.43-58.
- Kotimah, M.K. and Wulandari, S.P. (2014) 'Model Regresi Logistik Biner Stratifikasi Pada Partisipasi Ekonomi Perempuan Di Provinsi Jawa Timur', *Jurnal Sains dan Seni ITS*, 3(1), pp.D1-D6.
- Mahfudin, A. and Waqi'ah, K. (2016) 'Pernikahan Dini dan Pengaruhnya terhadap Keluarga di Kabupaten Sumenep Jawa Timur', *Jurnal Hukum Keluarga Islam*, 1(1), pp.33-49.
- Mariyatul Qibtiyah (2014) Faktor yang Mempengaruhi Perkawinan Muda Perempuan. 3(1):50-58
- Qibtiyah, M. (2015) 'Faktor yang mempengaruhi perkawinan muda perempuan. *Biometrika dan Kependudukan*', 3(1), pp.50-58.
- Purwa, T. (2019) 'Perbandingan Metode Regresi Logistik dan Random Forest untuk Klasifikasi Data Imbalanced (Studi Kasus: Klasifikasi Rumah Tangga Miskin di Kabupaten Karangasem, Bali Tahun 2017)', *Jurnal Matematika, Statistika dan Komputasi*, 16(1), pp.58-73.
- Salamah, S. (2016) *Faktor-faktor yang berhubungan dengan pernikahan usia dini di kecamatan pulokulon kabupaten grobogan*. Doctoral dissertation, Universitas Negeri Semarang.
- Verdikha, N.A., Adji, T.B. and Permanasari, A.E. (2018) 'Komparasi Metode Oversampling untuk Klasifikasi Teks Ujaran Kebencian'. *SEMNASTEKNO MEDIA ONLINE*, 6(1), pp.1-2.
- Xiao, J., Xie, L., He, C. and Jiang, X. (2012) 'Dynamic classifier ensemble model for customer classification with imbalanced class distribution', *Expert Systems with Applications*, 39(3), pp.3668-3675.