



TUGAS AKHIR - KS184822

**KLASIFIKASI KEJADIAN KEJANG PENDERITA EPILEPSI
BERDASARKAN FAKTOR RISIKO KEJANG EPILEPSI
MENGUNAKAN REGRESI LOGISTIK BINER DAN
NAÏVE BAYES**

**LIANNA DWI RAHMAWATI
NRP 062115 4000 0055**

**Dosen Pembimbing
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



TUGAS AKHIR - KS184822

**KLASIFIKASI KEJADIAN KEJANG PENDERITA EPILEPSI
BERDASARKAN FAKTOR RISIKO KEJANG EPILEPSI
MENGUNAKAN REGRESI LOGISTIK BINER DAN
NAÏVE BAYES**

**LIANNA DWI RAHMAWATI
NRP 062115 4000 0055**

**Dosen Pembimbing
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



FINAL PROJECT - KS184822

**CLASSIFICATION OF SEIZURE OF EPILEPSY PATIENTS
BASED ON RISK FACTOR USING BINARY LOGISTIC
REGRESSION AND NAÏVE BAYES**

**LIANNA DWI RAHMAWATI
SN 062115 4000 0055**

**Supervisor
Dr. Santi Wulan Purnami, S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

LEMBAR PENGESAHAN

KLASIFIKASI KEJADIAN KEJANG PENDERITA EPILEPSI BERDASARKAN FAKTOR RISIKO KEJANG EPILEPSI MENGUNAKAN REGRESI LOGISTIK BINER DAN NAÏVE BAYES

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember


Oleh :

Lianna Dwi Rahmawati
NRP. 062115 4000 0055

Disetujui oleh Pembimbing :
Dr. Santi Wulan Purnami, S.Si., M.Si.
NIP. 19720923 199803 2 001



Mengetahui,
Kepala Departemen Statistika


Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

KLASIFIKASI KEJADIAN KEJANG PENDERITA EPILEPSI BERDASARKAN FAKTOR RISIKO KEJANG EPILEPSI MENGGUNAKAN REGRESI LOGISTIK BINER DAN NAÏVE BAYES

Nama Mahasiswa : Lianna Dwi Rahmawati
NRP : 062115 4000 0055
Departemen : Statistika-FMKSD-ITS
Dosen Pembimbing : Dr. Santi Wulan Purnami, S.Si., M.Si.

Abstrak

Epilepsi merupakan gangguan susunan saraf pusat yang dicirikan oleh terjadinya serangan kejang secara spontan dan berkala. Kejang dapat diartikan sebagai modifikasi fungsi otak yang bersifat mendadak dan sepiantas yang berasal dari sekelompok sel-sel otak. Banyak hal yang dapat memicu kambuhnya kejang pada penderita epilepsi, seperti kurang tidur, stres, konsumsi alkohol, dan stimulasi seperti cahaya televisi atau lampu disko. Mayoritas dari penderita epilepsi hanya mengetahui satu atau dua hal yang menjadi pemicu terjadinya kejang. Akibatnya kejadian kejang yang dialami tidak terkontrol. Oleh karena itu, diperlukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi kejadian kejang dan prediksi kejadian kejang pada penderita epilepsi agar penderita epilepsi dapat mengantisipasi kemungkinan kejang yang akan terjadi. Penelitian ini menggunakan data dari penderita epilepsi yang tergabung dalam Komunitas ODE. Metode analisis yang digunakan adalah regresi logistik biner dan naïve bayes. Pada analisis regresi logistik biner menghasilkan empat faktor yang signifikan berpengaruh terhadap kejadian kejang pada penderita epilepsi yaitu usia penderita, usia saat terdiagnosis epilepsi, riwayat demam disertai kejang, dan faktor pemicu kejang. Hasil prediksi kejadian kejang menggunakan metode regresi logistik biner memiliki ketepatan prediksi sebesar 75%, sedangkan metode naïve bayes memiliki ketepatan prediksi yang lebih kecil yaitu sebesar 66,67%.

Kata Kunci: *Epilepsi, Faktor Risiko, Kejadian Kejang, Naïve Bayes, Regresi Logistik Biner*

(Halaman ini sengaja dikosongkan)

CLASSIFICATION OF SEIZURE OF EPILEPSY PATIENTS BASED ON RISK FACTOR USING BINARY LOGISTIC REGRESSION AND NAÏVE BAYES

Name : Lianna Dwi Rahmawati
Student Number : 062115 4000 0055
Department : Statistics
Supervisor : Dr. Santi Wulan Purnami, S.Si., M.Si.

Abstract

Epilepsy is a central nervous system disorder characterized by periodic and spontaneous seizures. Seizures can be interpreted as a sudden and fleeting modification of brain function that comes from a group of brain cells. Many things can trigger seizures, such as lack of sleep, stress, alcohol consumption, and stimulation such as television light or disco lights. Most epilepsy patient only know one or two things that trigger the event of seizures. As a result the event of seizures is not controlled. Therefore, research is needed to determine the factors that influence the event of seizures and predict the event of seizures of epilepsy patients, so that epilepsy patients can anticipate the possibility of seizures that will occur. This research uses data from epilepsy patients who are members of the ODE Community. The analytical method used is binary logistic regression and naïve bayes classification. The result of binary logistic regression analysis are four factors that have a significant effect on the event of seizures of epilepsy patients including the age of the epilepsy patient, age diagnosed with epilepsy, history of fever with seizures, and seizure trigger factors. The prediction results of seizure events using the binary logistic regression have predictive accuracy of 75%, while the naïve bayes method has a smaller predictive accuracy of 66,67%

Keywords: *Binary Logistic Regression, Epilepsy, Incident of Seizures, Naïve Bayes, Risk Factor*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat, taufiq serta hidayah-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Klasifikasi Kejadian Kejang Penderita Epilepsi Berdasarkan Faktor Risiko Kejang Epilepsi Menggunakan Regresi Logistik Biner dan Naïve Bayes”**. Sholawat serta salam semoga tetap tercurahkan kepada junjungan Nabi Muhammad SAW beserta keluarga dan para sahabat. Penyusunan Tugas Akhir ini dapat terselesaikan dengan baik dan lancar karena tidak lepas dari dukungan berbagai pihak. Oleh karena itu, penulis mengucapkan terimakasih kepada:

1. Bapak dan Ibu tersayang, Luhur Budi Santoso dan Tri Rahajoeningtyas yang selama ini telah memberikan segalanya bagi penulis, do'a, cinta, kasih sayang, pengorbanan dan perjuangan sehingga penulis bisa sampai seperti sekarang ini.
2. Dr. Suhartono selaku Kepala Departemen Statistika ITS dan sekaligus selaku dosen wali dari semester 1 sampai semester 8 yang telah membimbing, memberikan nasehat serta motivasi kepada penulis selama menempuh pendidikan.
3. Dr. Santi Wulan Purnami, S.Si., M.Si. selaku Kepala Program Studi Sarjana Departemen Statistika ITS dan sekaligus selaku dosen pembimbing yang telah memberikan banyak bantuan, masukan, arahan dan juga dukungan bagi penulis untuk dapat menyelesaikan Tugas Akhir ini.
4. Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D. dan Dr. Sutikno, S.Si., M.Si. selaku dosen penguji yang telah memberikan masukan berupa kritik dan saran untuk kesempurnaan Tugas Akhir ini.
5. Keluarga besar Departemen Statistika ITS, segenap dosen pengajar, staf dan karyawan yang telah memberikan ilmu selama penulis menempuh pendidikan dan telah membantu kelancaran dalam pelaksanaan kegiatan perkuliahan.
6. Dwinanto, Yoesi Irwatingrum, Talitha Prima Islamey, Chairunissa Kusnoaji, dan seluruh keluarga besar Almarhum Soenarto yang selalu memberikan do'a serta dukungan baik

secara moril maupun materil untuk penulis sehingga penulis dilancarkan dalam menyelesaikan Tugas Akhir ini.

7. Anadia Rahmat Syihab Hidayatullah yang selalu sabar menemani, memberikan dukungan, dan motivasi kepada penulis selama menjalani masa perkuliahan sampai penyusunan laporan Tugas Akhir.
8. Fraschiska Rizky Restuningtyas teman satu project dengan penulis yang telah bekerjasama dalam pengumpulan data dan saling memberikan semangat dalam penyusunan Tugas Akhir.
9. Shindy Sari Utami, Rhavida Anniza, dan Fonda Leviany yang telah banyak membantu dan mendukung penulis dalam menyelesaikan Tugas Akhir.
10. Teman-teman Vivacious Σ26 yang telah menjadi keluarga semenjak penulis menempuh perkuliahan di Departemen Statistika ITS.
11. Ibu Delyma dan anggota Komunitas ODE yang telah membantu dalam pengumpulan data dan berkenan untuk menjadi responden dalam penelitian Tugas Akhir ini.
12. Semua pihak yang telah memberikan dukungan dan bantuan dalam keberhasilan proses panjang ini yang tidak dapat disebutkan satu persatu namanya oleh penulis.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari kata sempurna. Oleh karena itu penulis sangat mengharapkan kritik dan saran yang membangun agar berguna untuk perbaikan berikutnya. Semoga laporan Tugas Akhir ini bermanfaat.

Surabaya, Mei 2019

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	5
1.4 Manfaat	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif	7
2.2 Tabel Kontingensi dan Uji Independensi	7
2.3 Regresi Logistik Biner	8
2.4 Naïve Bayes Classifier	15
2.5 Evaluasi Klasifikasi Model	16
2.6 Kejang Epilepsi dan Faktor-Faktor Risikonya	18
BAB III METODOLOGI PENELITIAN	21
3.1 Sumber Data	21
3.2 Teknik Pengambilan Sampel	21
3.3 Kerangka Konsep	22
3.4 Instrumen Penelitian	24
3.5 Variabel Penelitian dan Definisi Operasional	25
3.6 Struktur Data Penelitian	27
3.7 Langkah Penelitian	27

BAB IV ANALISIS DAN PEMBAHASAN	33
4.1 Karakteristik Penderita Epilepsi.....	33
4.2 Analisis Regresi Logistik Biner Berdasarkan Faktor- Faktor yang Mempengaruhi Kejadian Kejang	40
4.2.1 Uji Independensi Antara Kejadian Kejang dengan Faktor-Faktor Risiko Kejadian Kejang	41
4.2.2 Pembentukan Model Regresi Logistik Biner dengan Seluruh Variabel.....	42
4.2.3 Pembentukan Model Regresi Logistik Biner dengan <i>Backward Elimination</i>	44
4.3 Klasifikasi Kejadian Kejang Penderita Epilepsi dengan Naïve Bayes	48
BAB V KESIMPULAN DAN SARAN	59
5.1 Kesimpulan	59
5.2 Saran.....	60
DAFTAR PUSTAKA.....	61
LAMPIRAN.....	65

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Kurva ROC Sumber: (James dkk., 2013).....	18
Gambar 3.1 Ilustrasi Pengambilan Sampel Penelitian	22
Gambar 3.2 Kerangka Konsep Penelitian	23
Gambar 3.3 Diagram Alir Penelitian	29
Gambar 3.4 Diagram Alir Klasifikasi Menggunakan Regresi Logistik Biner	30
Gambar 3.5 Diagram Alir Klasifikasi dengan Naïve Bayes.....	31
Gambar 4.1 Proporsi Penderita Epilepsi Berdasarkan Pendidikan Terakhir	33
Gambar 4.2 Jumlah Penderita Epilepsi Berdasarkan Frekuensi Kejang	34
Gambar 4.3 Proporsi Penderita Epilepsi pada Tiap Kategori Kejadian Kejang	35
Gambar 4.4 Proporsi Penderita Epilepsi Berdasarkan Jenis Kelamin Antar Kategori Kejadian Kejang	36
Gambar 4.5 Perbedaan Median Usia Penderita Epilepsi Antar Kategori Kejadian Kejang	37
Gambar 4.6 Perbedaan Median Usia Terdiagnosis Penderita Epilepsi Antar Kategori Kejadian Kejang	37
Gambar 4.7 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Riwayat Demam Disertai Kejang	38
Gambar 4.8 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Kepatuhan Konsumsi Obat.....	38
Gambar 4.9 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Pemicu Kejang.....	39
Gambar 4.10 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Intensitas Terbangun	39
Gambar 4.11 Perbandingan Ketepatan Klasifikasi Tiap <i>Fold</i> pada Data <i>Testing</i> dengan Metode Regresi Logistik Biner	41

Gambar 4.12 Perbandingan Ketepatan Klasifikasi Tiap *Fold*
pada Data *Testing* dengan Metode Naïve Bayes 49

DAFTAR TABEL

	Halaman
Tabel 2.1 Tabel Kontingensi Dua Dimensi	7
Tabel 2.2 Nilai Regresi Logistik	13
Tabel 2.3 <i>Confusion Matrix</i> Dua Kelas	17
Tabel 2.4 Kriteria Nilai AUC	18
Tabel 3.1 Hasil <i>Purposive Sampling</i>	21
Tabel 3.2 Aspek dan Pertanyaan Penelitian.....	24
Tabel 3.3 Variabel Penelitian.....	25
Tabel 3.4 Struktur Data Penelitian	27
Tabel 4.1 Nilai AUC, Sensitivitas, dan Spesifisitas Tiap <i>Fold</i> pada Data <i>Testing</i>	40
Tabel 4.2 Hasil Uji Independensi Antara Kejadian Kejang dengan Faktor Risiko	41
Tabel 4.3 Hasil Uji Serentak Regresi Logistik Biner Tahap Awal.....	43
Tabel 4.4 Hasil Uji Parsial Regresi Logistik Biner Tahap Awal.....	43
Tabel 4.5 Hasil Uji Serentak Setelah Dilakukan <i>Backward</i> <i>Elimination</i>	45
Tabel 4.6 Hasil Uji Parsial Setelah Dilakukan <i>Backward</i> <i>Elimination</i>	45
Tabel 4.7 <i>Confusion Matrix</i> Regresi Logistik Biner Hasil Seleksi <i>Backward</i>	47
Tabel 4.8 Nilai AUC, Sensitivitas, dan Spesifisitas Tiap <i>Fold</i> pada Data <i>Testing</i>	49
Tabel 4.9 Nilai Probabilitas Awal pada Data <i>Training</i>	50
Tabel 4.10 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Jenis Kelamin	51
Tabel 4.11 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Riwayat Demam Kejang	51
Tabel 4.12 Nilai Probabilitas Bersyarat Variabel Kejadian	

	Kejang Berdasarkan Konsumsi Obat Antiepilepsi.....	52
Tabel 4.13	Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Pemicu Kejang.....	53
Tabel 4.14	Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Intensitas Terbangun	54
Tabel 4.15	Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Pemakaian AC	54
Tabel 4.16	Nilai Rata-Rata dan Standar Deviasi Variabel Usia dan Usia Terdiagnosis Tiap Kategori Kejadian Kejang	55
Tabel 4.17	Probabilitas Bersyarat Tiap Kategori pada Data <i>Testing</i> Pertama	55
Tabel 4.18	<i>Confusion Matrix</i> Klasifikasi Naïve Bayes	57

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Kuisisioner Penelitian.....	65
Lampiran 2. Data Pengujian.....	71
Lampiran 3. <i>Syntax k-Fold Cross Validation</i> Regresi Logistik Biner.....	72
Lampiran 4. <i>Syntax k-Fold Cross Validation</i> Naïve Bayes <i>Classification</i>	73
Lampiran 5. Tabel Kontingensi Antara Variabel Kejadian Kejang dengan Variabel Prediktor yang Bersifat Kategorik.....	74
Lampiran 6. <i>Output</i> Regresi Logistik Biner Tahap Awal	78
Lampiran 7. Pemilihan Model Terbaik Menggunakan <i>Backward Elimination</i>	79
Lampiran 8. Analisis Regresi Logistik Biner pada Model Terbaik	81
Lampiran 9. <i>Confusion Matrix</i> Data <i>Testing</i> untuk Regresi Logistik Biner	82
Lampiran 10. <i>Output</i> Naïve Bayes dari Data <i>Training</i>	83
Lampiran 11. <i>Confusion Matrix</i> Data <i>Testing</i> untuk Naïve Bayes.....	69
Lampiran 12. Surat Pernyataan Keaslian Data	85

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Epilepsi berasal dari bahasa Yunani, *Epilambanmein*, yang berarti serangan (Tunnajah, 2005). Dahulu epilepsi dikenal oleh masyarakat sebagai penyakit yang disebabkan oleh roh jahat dan dipercaya juga bahwa epilepsi merupakan penyakit kutukan. Kepercayaan tersebut menyebabkan munculnya stigma negatif bagi masyarakat. Masyarakat cenderung untuk menjauhi para penderita epilepsi karena epilepsi dianggap sebagai penyakit menular (melalui air liur yang keluar dari mulut), penyakit keturunan, menakutkan, memalukan, dan tidak bisa disembuhkan (Hauser, 1994). Mitos mengenai penyakit epilepsi terus berkembang di masyarakat sampai orang pertama, Hipokrates, berhasil mengenal epilepsi sebagai gejala penyakit dan menganggap bahwa epilepsi merupakan penyakit yang didasari oleh adanya gangguan di otak. Adelow dkk (2006) juga mendefinisikan epilepsi sebagai salah satu penyakit yang banyak dikenal manusia dan merupakan gangguan neurologis umum yang mempengaruhi sistem saraf pusat yang dapat terjadi pada setiap orang di seluruh dunia.

Epilepsi dapat dijumpai pada laki-laki maupun perempuan, tanpa memandang usia dan ras dengan insidensi dan prevalensi yang hampir sama walaupun beberapa peneliti menemukan angka lebih tinggi di negara berkembang. Selain itu banyak juga ditemukan bahwa penderita laki-laki umumnya sedikit lebih banyak daripada penderita perempuan, dan lebih sering dijumpai pada anak pertama (Tjahjadi, Dikot, dan Gunawan, 2005). Secara umum diperoleh gambaran bahwa insiden epilepsi menunjukkan pola bimodal, dimana puncak insiden terdapat pada golongan anak dan lanjut usia. Gambaran tersebut sejalan dengan penelitian Soetomenggolo dan Ismael (1999) yang menyatakan bahwa epilepsi merupakan masalah pediatrik yang besar dan lebih sering terjadi pada usia dini dibandingkan usia selanjutnya. Penelitian insidensi dan prevalensi telah dilaporkan oleh berbagai negara, tetapi di Indonesia belum diketahui secara pasti. Para peneliti umumnya mendapatkan insiden yang mencapai 50 per 100.000 per tahun dan prevalensi sekitar 6-10

per 1000 pada populasi umum. Sehingga dengan jumlah penduduk Indonesia yang mencapai 250 juta jiwa diperkirakan terdapat 2,5 juta jiwa yang menderita epilepsi pada saat ini (Sompa, 2016).

Menurut Harsono (1999), tidak ada penyebab tunggal pada epilepsi karena banyak faktor yang dapat menciderai sel-sel saraf otak. Penyakit epilepsi yang diketahui penyebabnya disebut epilepsi simtomatik, dan beberapa faktor risiko yang sudah diketahui antara lain: trauma kepala, demam tinggi, stroke, intoksikasi (termasuk obat-obatan), tumor otak, masalah kardiovaskuler tertentu, gangguan keseimbangan elektrolit, penyakit meningitis, dan infeksi parasit terutama cacing pita. Meskipun telah diketahui beberapa penyebab penyakit epilepsi, namun lebih dari 65% dari seluruh kasus epilepsi tidak diketahui faktor penyebabnya. Epilepsi yang tidak diketahui penyebabnya disebut epilepsi idiopatik (Raharjo, 2007). Epilepsi merupakan salah satu masalah kesehatan yang menjadi problem medis sekaligus menjadi problem sosial. Menjadi problem medis karena epilepsi membutuhkan penanganan dan pengawasan yang ketat dalam pengobatannya, sedangkan permasalahan sosial yang dihadapi penderita epilepsi akan lebih berat dibandingkan permasalahan medis yang dialaminya. Permasalahan sosial tersebut mencakup masalah psikososial yang dialami oleh penderita epilepsi, dimana adanya perasaan takut bahwa sepanjang hidupnya akan menderita epilepsi, takut mengemudi, takut berenang, dan yang paling ditakuti yaitu jika terjadi serangan kejang di depan umum (Salsabila, 2012).

Menurut Riyadi dan Sukarmin (2009) epilepsi merupakan gangguan susunan saraf pusat yang dicirikan oleh terjadinya serangan yang bersifat spontan dan berkala. Serangan dapat diartikan sebagai modifikasi fungsi otak yang bersifat mendadak dan sepiintas yang berasal dari sekelompok sel-sel otak, bersifat sinkron dan berirama. Kejang merupakan ciri epilepsi yang harus ada, tetapi tidak semua kejang merupakan manifestasi epilepsi. Kejang berulang yang terjadi pada penderita epilepsi terjadi akibat lepasnya muatan listrik neuron otak secara berlebihan dan paroksimal (Engel dan Pedley, 2008). Terdapat dua kategori dari kejang epilepsi yaitu kejang fokal (parsial) dan kejang umum. Kejang parsial terjadi karena adanya lesi pada satu bagian dari *cerebral cortex*, dimana pada kelainan ini dapat

disertai kehilangan kesadaran parsial. Sedangkan pada kejang umum, lesi mencakup area yang luas dari *cerebral cortex* dan biasanya mengenai kedua *hemisfer cerebri*. Kejang mioklonik, atonik, dan tonik-klonik termasuk dalam jenis kejang umum, dan dari beberapa jenis kejang yang ada kejang tonik-klonik adalah kejang yang sering ditemui. Kejang jenis ini biasanya berlangsung sekitar dua menit atau kurang, dan frekuensi serangannya dapat terjadi setiap jam sampai satu tahun sekali. Kambuhnya kejang pada seseorang yang diketahui mengidap epilepsi, bisa diprovokasi oleh banyak hal seperti kurang tidur, stres, alkohol, dan kadang-kadang stimulasi seperti cahaya televisi atau lampu disko. Pada wanita yang sedang menstruasi frekuensi kejang yang dialami juga bisa meningkat (Rubeinstain, 2007). Namun kebanyakan dari penderita epilepsi hanya mengetahui satu atau dua hal saja yang menjadi pemicu terjadinya serangan kejang, tanpa memperhatikan faktor-faktor dari aspek lainnya.

Penelitian mengenai kejadian serangan kejang pada penderita epilepsi masih jarang dan bersifat terbatas pada satu atau dua faktor penyebab saja, dan kebanyakan masih bersifat deskriptif. Berdasarkan hal tersebut maka perlu dilakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi kejadian serangan kejang pada penderita epilepsi. Menurut Kristanto (2017) faktor penyebab timbulnya kejang antara lain ketidakpatuhan minum obat, konsumsi alkohol, kurang tidur, dan konsumsi obat lain yang dapat mengganggu efek obat epilepsi. Penelitian mengenai faktor risiko kejang epilepsi telah dilakukan oleh Rakers dkk (2017) yang menganalisis hubungan kondisi cuaca dengan kejang pada penderita epilepsi menggunakan *odds ratio* dan memberikan kesimpulan bahwa tekanan atmosfer yang rendah dan kelembaban udara yang relatif tinggi dapat meningkatkan risiko kejadian kejang epilepsi, sedangkan cuaca dengan suhu yang tinggi dapat menurunkan risiko kejang. Penelitian lain tentang faktor risiko kejang juga pernah dilakukan oleh Nisa (2015), yaitu penelitian mengenai hubungan kualitas tidur terhadap kejadian kejang pada penderita epilepsi. Penelitian Tripathi dkk (2013) di India melaporkan bahwa putus obat adalah sebagai faktor utama pencetus bangkitan epilepsi yakni sebanyak 40,9% dari kasus yang ada. Ketiga penelitian tersebut dilakukan hanya untuk mengetahui ada atau tidaknya hubungan

antara kejadian kejang dengan beberapa variabel yang diduga mempengaruhi. Oleh karena itu, diperlukan penelitian untuk mengetahui faktor-faktor yang memengaruhi kejadian kejang dan prediksi kejadian kejang pada penderita epilepsi agar penderita epilepsi dapat mengantisipasi kemungkinan kejang yang akan terjadi.

Salah satu metode statistika yang dapat digunakan untuk mengklasifikasikan kejadian kejang adalah regresi logistik biner. Regresi logistik biner adalah suatu metode yang dapat digunakan untuk klasifikasi suatu data berdasarkan model yang didapatkan dari variabel respon yang bersifat biner dengan variabel prediktor yang berupa data kategorik maupun kontinyu. Metode regresi logistik biner adalah salah satu dari sekian banyak metode yang dapat digunakan untuk menyelesaikan berbagai permasalahan termasuk klasifikasi (Samosir, 2015). Metode lain yang dapat digunakan untuk klasifikasi adalah *naïve bayes classification*, dimana metode pengklasifikasiannya adalah berdasarkan probabilitas sederhana dan dirancang untuk dipergunakan dengan asumsi bahwa antar satu kelas dengan kelas yang lain saling independen.

Pada penelitian ini akan dilakukan klasifikasi kejadian kejang berdasarkan faktor-faktor risiko kejang pada penderita epilepsi dengan menggunakan sampel penderita epilepsi dalam Komunitas ODE. Metode yang digunakan adalah regresi logistik biner dan *naïve bayes*. Kedua metode tersebut memiliki kelebihan masing-masing. Kelebihan dari regresi logistik menurut Ghozali (2009) adalah analisis regresi logistik tidak memerlukan asumsi normalitas pada variabel bebasnya. Artinya, variabel penjelasnya tidak harus memiliki distribusi normal, linear, maupun memiliki varian yang sama dalam setiap kelompok. Gujarati (2003) menyatakan bahwa regresi logistik juga mengabaikan masalah heteroskedastisitas, sehingga tahapan analisis regresi logistik biner akan terdiri dari penjelasan statistika deskriptif dan pengujian hipotesis penelitian saja tanpa adanya pengecekan asumsi. Sedangkan kelebihan metode *naïve bayes* adalah memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya (Xhemali, Hinde, dan Stone, 2009). Selain itu kelebihan lain dari *naïve bayes* adalah hanya membutuhkan data *training* yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Beberapa

penelitian yang dilakukan untuk membandingkan akurasi yang dihasilkan oleh kedua metode tersebut memberikan hasil bahwa performansi dari model yang terbentuk dari regresi logistik biner lebih baik. Seperti pada penelitian yang dilakukan oleh Rajagukguk (2015) yang melakukan perbandingan metode klasifikasi regresi logistik biner dan naïve bayes pada status pengguna KB di Kota Tegal, dimana penelitian tersebut memberikan hasil ketepatan klasifikasi dengan metode regresi logistik biner sebesar 83,33% yang lebih besar dari ketepatan klasifikasi dengan metode naïve bayes yaitu sebesar 81,75%. Berdasarkan uraian tersebut maka pada penelitian ini digunakan regresi logistik biner dan naïve bayes untuk mengklasifikasikan kejadian kejang penderita epilepsi dalam Komunitas ODE.

1.2 Rumusan Masalah

Epilepsi merupakan gangguan sistem saraf pusat yang dicirikan dengan sering terjadinya kejang pada penderitanya. Banyak hal yang dapat memicu kambuhnya kejang. Namun, mayoritas dari penderita epilepsi hanya menyadari satu atau dua hal saja yang dapat memicu terjadinya kejang. Akibatnya kejang yang dialami penderita sering tidak terkontrol. Berdasarkan hal tersebut perlu diketahui faktor-faktor yang mempengaruhi kejadian kejang pada penderita epilepsi dan prediksi kejadian kejang berdasarkan faktor-faktor yang berpengaruh agar penderita epilepsi dapat mengantisipasi kemungkinan kejang yang akan terjadi.

1.3 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan, tujuan yang ingin dicapai dalam penelitian ini yaitu

1. Mendeskripsikan karakteristik penderita epilepsi dalam Komunitas ODE.
2. Mengetahui faktor-faktor yang mempengaruhi kejadian kejang pada penderita epilepsi dan memprediksi kejadian kejang pada penderita epilepsi dengan metode regresi logistik biner dan naïve bayes.

1.4 Manfaat

Hasil dari penelitian ini dapat bermanfaat bagi penderita epilepsi untuk mengetahui segala faktor yang dapat mempengaruhi

kejadian kejang epilepsi. Hasil pengklasifikasian kejadian kejang diharapkan dapat menambah informasi dan pengetahuan bagi tenaga kesehatan dalam rangka memberikan pelayanan kesehatan kepada penderita dengan tatalaksana yang lebih tepat. Selain itu hasil dari pemodelan kejadian kejang diharapkan pula dapat membantu untuk melakukan pencegahan sedini mungkin agar tidak muncul kasus-kasus yang baru pada penyakit epilepsi yang disebabkan oleh kejang yang tidak terkontrol.

1.5 Batasan Masalah

Penelitian ini hanya berfokus pada penderita epilepsi yang menjadi unit penelitian, yaitu penderita epilepsi dalam Komunitas ODE. Kejadian kejang dibedakan menjadi tidak mengalami kejang dan mengalami kejang, dimana kejang yang dimaksud adalah jenis kejang tonik-klonik yang dialami oleh penderita epilepsi dalam rentang waktu satu bulan. Metode yang digunakan untuk klasifikasi kejadian kejang penderita epilepsi dalam Komunitas ODE adalah regresi logistik biner dan naïve bayes. Ukuran ketepatan klasifikasi menggunakan nilai AUC, sensitivitas, dan spesifisitas.

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna (Walpole, 1995). Data yang disajikan dalam statistika deskriptif biasanya dalam bentuk ukuran pemusatan dan ukuran penyebaran data. Salah satu ukuran pemusatan data yang biasa digunakan adalah mean, sedangkan salah satu ukuran penyebaran data adalah varians. Cara penyajian data dalam statistika deskriptif dapat menggunakan tabel, diagram, maupun grafik. Statistika deskriptif hanya berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai suatu data atau keadaan atau fenomena, dengan kata lain hanya melihat gambaran secara umum dari data yang didapatkan.

2.2 Tabel Kontingensi dan Uji Independensi

Tabel kontingensi atau yang biasa disebut dengan *cross tabulation* merupakan teknik penyusunan data yang cukup sederhana untuk melihat hubungan antara dua atau lebih variabel secara simultan. Variabel yang dianalisis merupakan variabel kategorikal yang memiliki skala nominal atau skala ordinal. Secara umum, tabel kontingensi dapat disajikan dalam bentuk tabel $R \times C$ seperti Tabel 2.1.

Tabel 2.1 Tabel Kontingensi Dua Dimensi

Variabel x	Variabel y				Total
	1	2	...	C	
1	n_{11}	n_{12}	...	n_{1C}	n_{1+}
2	n_{21}	n_{23}	...	n_{2C}	n_{2+}
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
R	n_{R1}	n_{R2}	...	n_{RC}	n_{R+}
Total	n_{+1}	n_{+2}	...	n_{+C}	n_{++}

Keterangan:

n_{RC} : frekuensi pengamatan pada baris ke- R dan kolom ke- C

n_{R+} : total marjinal pada variabel baris

n_{+C} : total marjinal pada variabel kolom

n_{++} : total frekuensi pengamatan

Tabel kontingensi $R \times C$ terdiri atas dua variabel, yaitu variabel independen (X) dengan kategori yang dinyatakan dalam R baris dan variabel dependen (Y) dengan kategori yang dinyatakan dalam C kolom. Terdapat $R \times C$ sel dalam tabel yang berisi frekuensi pengamatan yang terjadi dari kombinasi kategori kedua variabel sehingga diperoleh data berkategori dalam bentuk kontingensi dua dimensi berukuran $R \times C$ (Agresti, 2013).

Komponen-komponen dalam tabel kontingensi tersebut selanjutnya digunakan untuk perhitungan statistik uji independensi. Uji independensi merupakan suatu uji yang digunakan untuk mengetahui ada atau tidaknya keterkaitan variabel-variabel yang diamati. Suatu variabel dikatakan independen apabila distribusi variabel tersebut tidak dipengaruhi oleh distribusi variabel lain. Uji yang sesuai untuk mengetahui hubungan antara dua variabel kategori yang berupa tabel kontingensi adalah dengan *Pearson chi-square test*. Hipotesis uji independensi adalah sebagai berikut:

H_0 : Tidak ada hubungan antara variabel x dan variabel y

H_1 : Ada hubungan antara variabel x dan variabel y

Statistik uji:

$$\chi^2 = \sum_{r=1}^R \sum_{c=1}^C \frac{(n_{rc} - e_{rc})^2}{e_{rc}} \quad (2.1)$$

dimana $e_{rc} = \frac{n_{r+} \times n_{+c}}{n_{++}}$

n_{rc} : frekuensi observasi ke- r dari variabel x dan ke- c dari variabel y

e_{rc} : frekuensi harapan ke- r dari variabel x dan ke- c dari variabel y

r : 1, 2, ..., R

c : 1, 2, ..., C .

Statistik uji χ^2 mengikuti distribusi *chi-square* dengan derajat bebas $db = (C - 1)(R - 1)$, sehingga diperoleh keputusan tolak H_0 jika nilai $\chi^2 > \chi^2_{(\alpha, db)}$ atau $p\text{-value} < \alpha$.

2.3 Regresi Logistik Biner

Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon yang bersifat biner dengan variabel prediktor yang bersifat kategorik maupun kontinyu (Hosmer dan Lemeshow, 2000). Variabel respon (y) terdiri dari dua kategori yaitu “sukses” dan “gagal” yang

dinotasikan dengan $y = 1$ untuk sukses dan $y = 0$ untuk gagal. Variabel y mengikuti Distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi adalah sebagai berikut:

$$f(y) = \pi^y (1 - \pi)^{1-y} \quad (2.2)$$

dengan $y = 0$ atau 1 . Jika $y = 0$ maka $f(y) = 1 - \pi$ dan jika $y = 1$ maka $f(y) = \pi$.

Pada regresi logistik dapat disusun model yang terdiri dari banyak variabel prediktor yang dikenal sebagai model multivariabel (Agresti, 2013). Model regresi logistik biner dengan variabel prediktor sebanyak p adalah sebagai berikut:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \quad (2.3)$$

Fungsi $\pi(x)$ adalah fungsi non linear sehingga perlu dilakukan transformasi logit untuk memperoleh fungsi linear yang dapat digunakan untuk melihat hubungan antara variabel respon (y) dengan variabel prediktor (x). Bentuk logit $\pi(x)$ apabila ditransformasikan akan menghasilkan fungsi logit $g(x)$ sebagai berikut:

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.4)$$

Selanjutnya model regresi logistik pada persamaan (2.4) dapat dituliskan dalam bentuk sebagai berikut:

$$\pi(x) = \frac{\exp[g(x)]}{1 + \exp[g(x)]} \quad (2.5)$$

untuk $i = 1, 2, \dots, n$ maka model regresi logistik dapat ditulis seperti berikut:

$$\pi(x_i) = \frac{\exp \left[\sum_{j=0}^p \beta_j x_{ij} \right]}{1 + \exp \left[\sum_{j=0}^p \beta_j x_{ij} \right]} \quad (2.6)$$

Estimasi Parameter

Metode yang paling umum digunakan untuk menduga parameter pada model regresi logistik adalah metode *Maximum*

Likelihood Estimation (MLE). MLE dapat dilakukan jika antara amatan satu dengan yang lain diasumsikan saling bebas. Metode ini dipilih karena mempunyai beberapa kelebihan dibandingkan dengan metode lain, diantaranya dapat digunakan untuk model tidak linier seperti regresi logistik, serta hasil penaksirannya *unbiased*. Nilai estimasi β didapatkan dengan memaksimumkan fungsi *likelihood* (Agresti, 2013). Pada regresi logistik biner, setiap pengamatan mengikuti Distribusi Bernoulli sehingga dapat ditentukan fungsi probabilitasnya seperti berikut:

$$f(y = y_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (2.7)$$

dimana $i = 1, 2, \dots, n$. Jika y_i dan x_i adalah pasangan variabel respon dan variabel prediktor pada pengamatan ke- i yang diasumsikan bahwa setiap pasangan pengamatan saling independen dengan pasangan pengamatan lainnya, maka persamaan *maximum likelihood* yang bisa didapatkan adalah sebagai berikut:

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n f(y_i) \\ &= \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \\ &= \left\{ \prod_{i=1}^n \exp \left[\ln \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right)^{y_i} \right] \right\} \left\{ \prod_{i=1}^n [1 - \pi(x_i)] \right\} \\ &= \left\{ \exp \left[\sum_{i=1}^n y_i \ln \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right) \right] \right\} \left\{ \prod_{i=1}^n [1 - \pi(x_i)] \right\} \quad (2.8) \\ &= \left\{ \exp \left[\sum_{i=1}^n y_i \ln \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right) \right] \right\} \left\{ \prod_{i=1}^n \frac{1}{1 + \exp \left(\sum_{j=0}^p \beta_j x_{ij} \right)} \right\} \\ &= \left\{ \exp \left(\sum_{i=1}^n y_i \sum_{j=0}^p \beta_j x_{ij} \right) \right\} \left\{ \prod_{i=1}^n \left[1 + \exp \left(\sum_{j=0}^p \beta_j x_{ij} \right) \right]^{-1} \right\} \end{aligned}$$

Fungsi *maximum likelihood* tersebut akan dimaksimumkan dalam bentuk $\ln L(\beta)$ dan dinyatakan dengan $l(\beta)$.

$$\begin{aligned}
 l(\beta) &= \ln[L(\beta)] \\
 &= \sum_{j=0}^p \left[\sum_{i=1}^n y_i x_{ij} \right] \beta_j - \sum_{i=1}^n \ln \left[1 + \exp \left(\sum_{j=0}^p \beta_j x_{ij} \right) \right]
 \end{aligned} \tag{2.9}$$

Nilai β maksimum didapatkan melalui turunan $l(\beta)$ terhadap β dan hasilnya adalah sama dengan nol.

$$\begin{aligned}
 \frac{\partial l(\beta)}{\partial \beta_j} &= 0 \\
 \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \left(\frac{\exp \left(\sum_{j=0}^p \beta_j x_{ij} \right)}{1 + \exp \left(\sum_{j=0}^p \beta_j x_{ij} \right)} \right) &= 0
 \end{aligned} \tag{2.10}$$

$$\sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \hat{\pi}(x_i) = 0$$

dimana $j = 0, 1, 2, \dots, p$. Pada saat dilakukan diferensial pertama dari fungsi *ln-likelihood* terhadap β ternyata memiliki hasil persamaan yang masih mengandung parameter lain atau dapat dikatakan bahwa persamaan hasil turunan pertama tidak linear, sehingga diperlukan iterasi numerik untuk memperoleh $\hat{\beta}$ yang konvergen. Metode iterasi *Newton Raphson* dapat digunakan untuk menyelesaikan persamaan non-linear. Dengan rumus iterasi seperti pada persamaan (2.11) (Agresti, 2013).

$$\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} - \left(\mathbf{H}(\hat{\beta}^{(t)}) \right)^{-1} \bar{g}^{(t)} ; t = 0, 1, 2, \dots \tag{2.11}$$

dimana

$$\begin{aligned}
 \bar{g} &= \left[\frac{\partial L(\beta)}{\partial \beta_0} \quad \frac{\partial L(\beta)}{\partial \beta_1} \quad \dots \quad \frac{\partial L(\beta)}{\partial \beta_p} \right]^T \\
 \mathbf{H} &= \begin{bmatrix} h_{11} & \dots & h_{1p} \\ \vdots & \ddots & \vdots \\ h_{p1} & \dots & h_{pp} \end{bmatrix} ; h_{a,b} = \frac{\partial^2 L(\beta)}{\partial \beta_a \partial \beta_b} ; a, b = 0, 1, 2, \dots
 \end{aligned}$$

Langkah-langkah iterasi *Newton Raphson* adalah sebagai berikut:

1. Menentukan nilai awal estimasi parameter $\hat{\beta}^{(0)}$

2. Membentuk vektor gradien \vec{g} dan matriks Hessian \mathbf{H}
3. Memasukkan nilai $\hat{\beta}^{(0)}$ pada elemen \vec{g} dan \mathbf{H} sehingga diperoleh $\vec{g}(\hat{\beta}^{(0)})$ dan $\mathbf{H}(\hat{\beta}^{(0)})$
4. Iterasi mulai $t = 0$ menggunakan persamaan (2.11). Nilai $\hat{\beta}^{(t)}$ merupakan sekumpulan penaksir parameter yang konvergen pada iterasi ke- t
5. Apabila belum diperoleh estimasi parameter yang konvergen, maka ulangi kembali langkah (3) hingga nilai $\|\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}\| \leq \varepsilon$, dengan ε merupakan bilangan yang sangat kecil. Hasil estimasi yang diperoleh adalah $\hat{\beta}^{(t+1)}$ pada iterasi terakhir.

Uji Signifikansi Parameter

Pengujian signifikansi parameter dilakukan untuk mengetahui berpengaruh atau tidaknya taksiran parameter terhadap model, dan juga untuk mengetahui besar pengaruh dari masing-masing parameter. Pengujian signifikansi parameter pada model regresi logistik biner dilakukan secara serentak dan parsial.

a. Uji Serentak

Pengujian signifikansi parameter secara serentak dilakukan untuk memeriksa kemaknaan koefisien β terhadap variabel respon secara bersama-sama. Salah satu statistik uji yang dapat digunakan yaitu *Likelihood Ratio Test*, dengan hipotesis

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

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0; j = 1, 2, \dots, p.$$

Statistik uji yang digunakan adalah statistik uji G atau *Likelihood Ratio Test* yang mengikuti distribusi *chi-square* (Hosmer dan Lemeshow, 2000)

$$G = -2 \ln \left[\frac{\binom{n_0}{n} \binom{n_1}{n}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{1-y_i}} \right] \quad (2.12)$$

Dimana,

$$n_0 = \sum_{i=1}^n y_{0i}, \quad n_1 = \sum_{i=1}^n y_{1i}, \quad n = n_0 + n_1.$$

Kriteria pengambilan keputusan menggunakan taraf nyata α . Jika $G > \chi^2_{(\alpha;p)}$ dengan p adalah banyaknya variabel prediktor, maka dapat diputuskan untuk menolak H_0 yang artinya minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon.

b. Uji Parsial

Pengujian parsial dilakukan guna mengetahui signifikansi variabel prediktor terhadap model. Menurut Kleinbaum dan Klein (2002), uji Wald dapat digunakan untuk menguji pengaruh dari satu parameter. Hipotesis yang digunakan dalam pengujian parsial yaitu $H_0 : \beta_j = 0$ (parameter dalam model, untuk variabel prediktor ke- k) dan hipotesis alternatifnya $H_1 : \beta_j \neq 0$, dimana $j = 1, 2, \dots, p$. Statistik uji Wald didapatkan dari pembagian antara parameter yang ditaksir terhadap parameter yang ditaksir oleh galat baku seperti pada persamaan (2.13).

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.13)$$

dimana $\hat{\beta}_j$ adalah penaksir parameter β_j dan $SE(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$. Statistik uji W mengikuti distribusi normal sehingga akan diperoleh keputusan tolak H_0 jika nilai $|W|$ lebih besar dari nilai $Z_{\alpha/2}$ atau p -value kurang dari α (Hosmer dan Lemeshow, 2000).

Interpretasi Koefisien Parameter

Interpretasi model dilakukan untuk mendefinisikan perubahan variabel respon yang disebabkan oleh variabel prediktor serta menentukan hubungan fungsional antara keduanya. Nilai *odds ratio* adalah salah satu ukuran yang dapat digunakan untuk menginterpretasikan koefisien parameter yang diperoleh dari nilai-nilai regresi logistik seperti yang ditunjukkan pada Tabel 2.2.

Tabel 2.2 Nilai Regresi Logistik

Variabel x	Variabel y	
	y = 1	y = 2
x = 1	$\pi_{11} = P(x = 1 y = 1)$	$\pi_{12} = P(x = 1 y = 2)$
x = 2	$\pi_{21} = P(x = 2 y = 1)$	$\pi_{22} = P(x = 2 y = 2)$

Berdasarkan Tabel 2.2 tersebut dapat diartikan bahwa *odds ratio* didefinisikan sebagai rasio nilai regresi logistik untuk $x = 1$ terhadap nilai regresi logistik untuk $x = 2$ (Agresti, 2013), yang dapat dirumuskan dalam persamaan berikut:

$$\Psi_r = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} \quad (2.14)$$

Nilai *odds ratio* selalu positif. Jika nilai $\Psi_r = 1$ menunjukkan kedua variabel tidak memiliki hubungan, sedangkan jika $0 < \Psi_r < 1$ maka variabel respon dan variabel prediktor memiliki hubungan negatif untuk setiap perubahan nilai pada variabel prediktor, dan jika nilai $\Psi_r > 1$ memiliki arti bahwa variabel respon berhubungan positif dengan variabel prediktor untuk setiap perubahan nilai pada variabel prediktor (Hosmer dan Lemeshow, 2000).

Uji Kesesuaian Model

Uji kesesuaian model digunakan untuk mengetahui apakah terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model. Hipotesis uji kesesuaian model adalah sebagai berikut:

H_0 : Model telah sesuai (Tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

H_1 : Model tidak sesuai (Terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

Statistik uji:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.15)$$

dimana,

O_k : nilai variabel respon pada grup ke- k ($\sum_{k=1}^{c_k} y_k$ dengan c_k : respon (0, 1, 2))

$\bar{\pi}_k$: rata-rata taksiran peluang ($\sum_{k=1}^{c_k} \frac{m_k \bar{\pi}_k}{n_k}$)

g : jumlah grup (kombinasi kategori dalam model serentak)

n'_k : banyak observasi pada grup ke- k .

Statistik uji \hat{C} mengikuti distribusi *chi-square* dengan derajat bebas $g - 2$, sehingga akan diperoleh keputusan tolak H_0 jika nilai \hat{C} lebih

besar dari nilai $\chi^2_{(\alpha, g-2)}$ atau p -value kurang dari α (Hosmer dan Lemeshow, 2000).

2.4 Naïve Bayes Classifier

Naïve Bayes Classifier (NBC) adalah salah satu algoritma dalam teknik data mining yang menerapkan teori Bayes dalam klasifikasi. Naïve bayes merupakan metode klasifikasi yang paling sering digunakan dan masuk dalam sepuluh algoritma terbaik dalam data mining (Wu dan Kumar, 2009). Jika diketahui x merupakan vector yang berisi fitur dan y adalah label kelas, maka naïve bayes dituliskan dengan $P(y | x)$. notasi tersebut berarti probabilitas label kelas y didapatkan setelah fitur-fitur x diamati. Notasi ini disebut juga probabilitas akhir (*posterior probability*) untuk y (Prasetyo, 2012). Formulasi naïve bayes untuk klasifikasi adalah sebagai berikut:

$$P(y | x_j) = \frac{P(y) \prod_{j=1}^p P(x_j | y)}{P(x_j)} \quad (2.16)$$

Dimana,

$P(y | x_j)$: probabilitas data dengan vektor x pada kelas y

$P(y)$: probabilitas awal kelas (*prior probability*)

$\prod_{j=1}^p P(x_j | y)$: probabilitas independen kelas y dari semua fitur dalam vektor x

$P(x_j)$: probabilitas dari x_j

Probabilitas $P(x_j)$ selalu tetap sehingga dalam perhitungan prediksi nantinya dapat dihilangkan dan hanya menghitung bagian $P(y) \prod_{j=1}^p P(x_j | y)$ saja dengan memilih nilai yang terbesar sebagai kelas yang dipilih sebagai hasil prediksi. Sementara probabilitas independen $\prod_{j=1}^p P(x_j | y)$ tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas y , yang dinotasikan dengan

$$P(x | y = a) = \prod_{j=1}^p P(x_j | y = a) \quad (2.17)$$

Untuk setiap set fitur $x = [x_1, x_2, \dots, x_p]$ terdiri atas p atribut. Jika terdapat variabel yang bersifat kontinyu maka $P(x_k | y_j)$ dihitung menggunakan pendekatan distribusi normal seperti berikut:

$$P(x_j | y) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{(x_j - \mu_j)^2}{2\sigma_j^2}\right) \quad (2.18)$$

Estimasi peluang $P(x_j|y)$ dapat dihitung untuk setiap variabel x_j dan kelas y sehingga data baru akan dapat diklasifikasikan ke dalam kelas y apabila peluangnya lebih besar dibandingkan yang lainnya. Pada naïve bayes digunakan *Hypothesis Maximum A Posterior* (HMAP) untuk memaksimalkan nilai probabilitas dari masing-masing kelas dengan rumus sebagai berikut: (Gorunescu, 2011)

$$\text{HMAP} = \arg \max \frac{P(x_1, x_2, \dots, x_p | y) \times P(y)}{P(x_1, x_2, \dots, x_p)} \quad (2.19)$$

2.5 Evaluasi Klasifikasi Model

Cross validation adalah bentuk sederhana dari teknik statistika untuk mengevaluasi kinerja model, *k-fold cross validation* adalah jenis *cross validation* yang digunakan pada kebanyakan penelitian karena dapat mengurangi waktu komputasi dengan tetap menjaga keakuratan estimasi. Dalam praktiknya, fold yang biasa digunakan yaitu *5-fold* dan *10-fold* (James dkk., 2013). Pada *k-fold cross validation* data yang digunakan akan dibagi secara acak ke dalam k subset yaitu D_1, D_2, \dots, D_k dengan ukuran yang sama. Dataset akan dibagi menjadi data *training* dan data *testing*. Proses *training* dan *testing* dilakukan sebanyak k kali secara berulang-ulang. Pada iterasi ke- i , partisi D_i disajikan sebagai data *testing* dan partisi sisanya digunakan secara bersamaan dan berurutan sebagai data *training*. Iterasi kedua, subset D_1, D_2, \dots, D_k akan dites pada D_2 . Proses tersebut terus berlangsung hingga dilakukan tes pada D_k (Han, Kamber, dan Pei, 2012).

Data *training* digunakan untuk membentuk model klasifikasi, yang selanjutnya model akan digunakan untuk mengklasifikasikan data *testing*. Ketepatan dari hasil klasifikasi dapat dihitung dengan menggunakan *confusion matrix* yang perhitungannya berdasarkan perkiraan objek yang benar atau salah. Sebuah matriks dari prediksi akan dibandingkan dengan kelas sebenarnya atau dengan kata lain berisi informasi nilai sebenarnya dan prediksi pada hasil klasifikasi

(Gorunescu, 2011). Komponen-komponen dalam *confusion matrix* adalah sebagai berikut:

- *True Positive* (TP), yaitu jumlah dokumen dari kelas 1 yang benar dan diklasifikasikan sebagai kelas 1
- *True Negative* (TN), yaitu jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0
- *False Positive* (FP), yaitu jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai 1
- *False Negative* (FN) yaitu jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0.

Agar lebih jelas *confusion matrix* disajikan pada Tabel 2.3.

Tabel 2.3 *Confusion Matrix* Dua Kelas

Kelas Sebenarnya	Kelas Prediksi	
	1	0
1	TP	FN
0	FP	TN

Setelah data *testing* diklasifikasikan maka akan didapatkan *confusion matrix* sehingga dapat dihitung nilai AUC, sensitivitas, dan spesifisitas dengan rumus sebagai berikut:

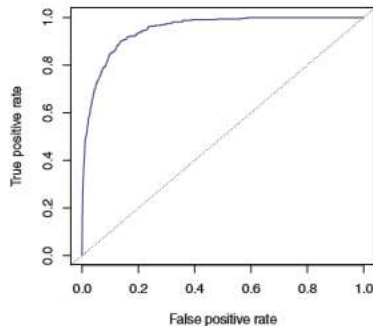
$$\text{AUC} = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (2.20)$$

$$\text{Sensitivitas} = \frac{TP}{TP + FN} \quad (2.21)$$

$$\text{Spesifisitas} = \frac{TN}{TN + FP} \quad (2.22)$$

AUC (*Area Under Curve* ROC) merupakan salah satu kriteria yang dapat digunakan untuk mengukur kualitas hasil dari metode klasifikasi (Huang dan Ling, 2005). AUC adalah area di bawah kurva *Receiver Operating Characteristic* (ROC). Kurva ROC adalah grafik populer untuk menampilkan secara bersamaan kurva dari dua jenis kesalahan untuk semua *threshold* yang mungkin. Kurva ROC yang ideal akan melengkung di sudut kiri atas seperti ditunjukkan pada Gambar 2.1, dimana garis lengkung menunjukkan tingkat *true positive* yang tinggi dan tingkat *false positive* yang rendah.

Sedangkan garis putus-putus menunjukkan tidak ada informasi yang didapatkan dari proses klasifikasi. Semakin besar area di bawah kurva ROC (AUC), maka semakin baik hasil klasifikasi yang dihasilkan (James dkk., 2013).



Gambar 2.1 Kurva ROC Sumber : (James dkk., 2013)

Kriteria keakuratan klasifikasi menggunakan AUC ditunjukkan dalam Tabel 2.4 (Suwarno dan Abdillah, 2016).

Tabel 2.4 Kriteria Nilai AUC

Nilai AUC	Interpretasi
0,9-1,0	<i>Excellent Classification</i>
0,8-0,9	<i>Good Classification</i>
0,7-0,8	<i>Fair Classification</i>
0,6-0,7	<i>Poor Classification</i>
0,5-0,6	<i>Failed</i>

Sedangkan untuk pengertian dari sensitivitas adalah proporsi dari kelas 1 yang teridentifikasi dengan benar, sedangkan spesifisitas adalah proporsi dari kelas 0 yang teridentifikasi dengan benar.

2.6 Kejang Epilepsi dan Faktor-Faktor Risikonya

Epilepsi adalah suatu kejadian lepasnya muatan listrik yang berlebihan dan abnormal dari sel-sel saraf otak secara berkala yang ditandai dengan kejang dan serangan yang berulang. Kejang adalah kejadian epilepsi dan merupakan ciri epilepsi yang harus ada, tetapi tidak semua kejang merupakan manifestasi epilepsi. Kejang terjadi akibat pengeluaran sejumlah neuron yang abnormal akibat dari

berbagai proses patologi sehingga berdampak pada otak. Secara umum dapat dikatakan bahwa serangan epilepsi dapat timbul jika terjadinya pelepasan aktivitas energi yang berlebihan dan mendadak dalam otak, sehingga menyebabkan terganggunya kerja otak (Harsono, 1999). Klasifikasi mengenai kejang yang sekarang dipergunakan secara luas adalah klasifikasi oleh *International League Against Epilepsy* (1997) yang terdiri dari 3 kategori utama yaitu kejang parsial, kejang umum dan kejang yang tak terklasifikasikan. Serangan epilepsi tidak selalu disertai dengan kejang dan sebaliknya, kejang belum tentu dapat dikatakan epilepsi. Salah satu serangan yang paling sering dijumpai pada penderita epilepsi yaitu serangan grandmal atau yang biasa disebut juga dengan serangan tonik-klonik atau bangkitan mayor atau *generalized tonic-clonic seizures* (GTCS). Serangan ini terjadi pada seluruh tubuh, dimulai dengan kakunya otot-otot tubuh yang kemudian diikuti oleh kontraksi otot-otot secara ritmik, dan dilanjutkan dengan kehilangan kesadaran.

Serangan kejang pada penderita epilepsi dipicu oleh beberapa faktor antara lain kurang tidur, ketidakpatuhan mengkonsumsi obat anti epilepsi, penyakit lain, perubahan hormon, stress, konsumsi alkohol atau obat-obatan, dehidrasi, dan stimulasi lampu kelap-kelip atau corak tertentu. Penelitian sebelumnya mengenai faktor-faktor yang mempengaruhi serangan kejang pada penderita epilepsi menurut Rubeinstain (2007) diprovokasi oleh

1. Kurang tidur
2. Stress
3. Konsumsi alkohol
4. Stimulasi seperti cahaya televisi atau lampu diskoteka
5. Waktu menstruasi pada perempuan

Penelitian serupa dilakukan oleh Silbernagl dan Lang (2006) tentang faktor-faktor yang mempengaruhi serangan kejang penderita epilepsi yaitu sebagai berikut:

1. Demam
2. Kurang tidur
3. Kelainan genetik
4. Konsumsi alkohol
5. Pendarahan

6. Pembengkakan sel

Penelitian lain tentang faktor-faktor yang mempengaruhi serangan kejang penderita epilepsi juga pernah dilakukan oleh Nisa (2015) dengan menggunakan kualitas tidur sebagai faktor yang mempengaruhi. Rakers dkk (2017) melakukan penelitian yang sama dengan faktor yang mempengaruhi serangan kejang penderita epilepsi yaitu tekanan atmosfer, suhu, dan kelembaban udara.

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data primer. Data primer yang digunakan mencakup informasi tentang gambaran kondisi responden meliputi riwayat kesehatan, sosio-ekonomi, dan gaya hidup responden selama menderita epilepsi yang diperoleh secara langsung dari penderita epilepsi yang menjadi unit penelitian. Adapun yang menjadi unit penelitian adalah 51 penderita epilepsi anak dan dewasa yang terdaftar dalam Komunitas ODE.

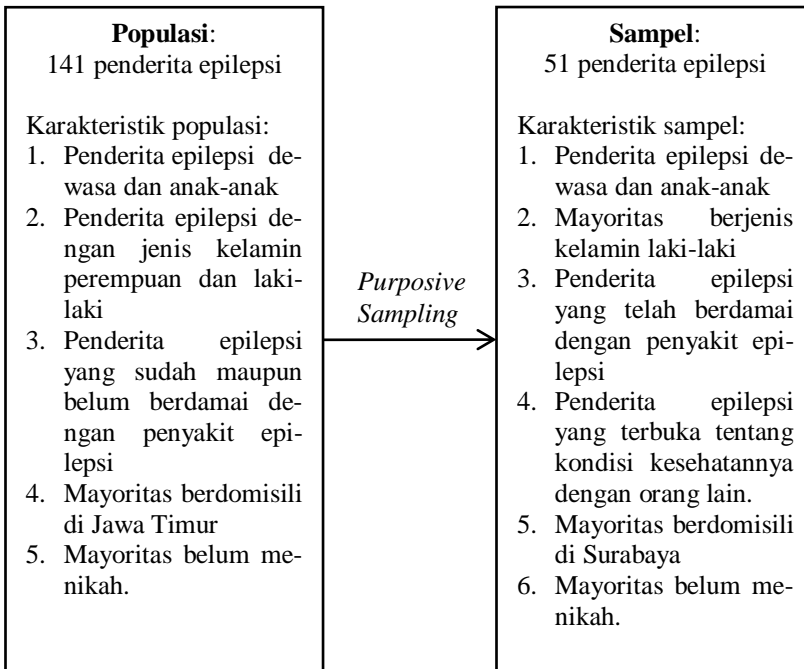
3.2 Teknik Pengambilan Sampel

Pada penelitian ini yang menjadi populasi adalah seluruh anggota Komunitas ODE yaitu sebanyak 141 penderita epilepsi. Penelitian terhadap populasi tidak dapat dilakukan karena tidak semua penderita epilepsi dalam Komunitas ODE dapat terbuka mengenai kondisi kesehatannya terhadap orang lain, sehingga perlu dilakukan penarikan sampel. Sampel dipilih menggunakan teknik *purposive sampling* dengan menetapkan kriteria-kriteria tertentu yang harus terpenuhi. Adapun kriteria yang harus terpenuhi untuk menjadi sampel penelitian yaitu penderita epilepsi dalam Komunitas ODE yang terbuka mengenai kondisi kesehatannya terhadap orang lain, dan bersedia menjadi responden. Berdasarkan kriteria yang telah ditetapkan, hasil dari *purposive sampling* ditunjukkan pada Tabel 3.1

Tabel 3.1 Hasil *Purposive Sampling*

Kriteria Sampel	Jumlah Sampel
Penderita epilepsi dalam Komunitas ODE yang terbuka mengenai kondisi kesehatannya kepada orang lain	54
Pengurangan Sampel:	
Penderita epilepsi yang tidak bersedia menjadi responden	(3)
Total Sampel	51

Berdasarkan Tabel 3.1 diketahui bahwa jumlah sampel yang terpilih yaitu sebanyak 51 sampel. Ilustrasi mengenai cara pengambilan sampel disajikan pada Gambar 3.1.



Gambar 3.1 Ilustrasi Pengambilan Sampel Penelitian

3.3 Kerangka Konsep

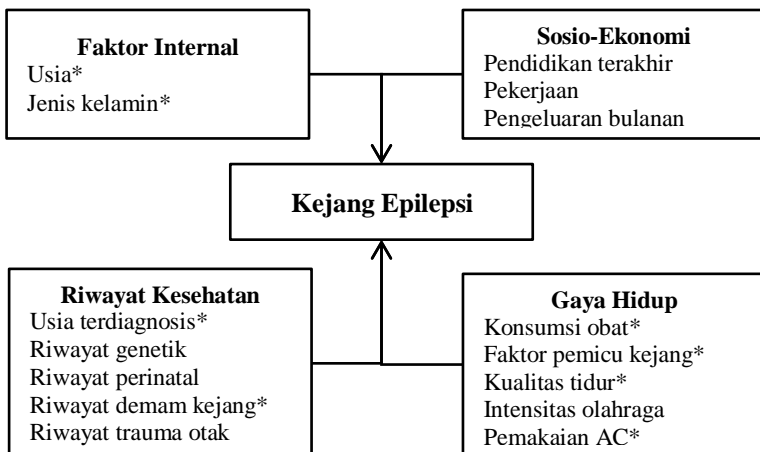
Faktor risiko yang dapat mempengaruhi peluang terjadinya kejadian kejang pada penderita epilepsi antara lain:

1. Jenis kelamin, menurut Yulinda (2015), menyebutkan bahwa epilepsi banyak diderita oleh laki-laki.
2. Usia pasien, disebutkan bahwa anak berusia dibawah 2 dan usia lanjut diatas 65 tahun akan lebih mungkin mengalami kejang (Vozikis, Goulionis, dan Nikolakis, 2012).
3. Penderita epilepsi dengan usia terdiagnosis kurang dari 3 tahun sering dikaitkan dengan berbagai sindrom epilepsi pada anak misalnya *Lennox Gestaut syndrome*, *West syndrome*, dan

Ohtahara syndrome yang akan memiliki frekuensi kejang lebih sering dan sulit diatasi (Sari, Triono, dan Sutomo, 2017)

4. Kejang demam, menurut Lynette dan Ingrid (2007) kejang demam merupakan salah satu faktor risiko paling umum yang mendasari epilepsi pada anak. Pada umumnya kasus kejang demam dapat sembuh sempurna, dan sebagian kecil berkembang menjadi kejang dari penyakit epilepsi (Knudzen, 2000)
5. Konsumsi obat anti epilepsi, disebutkan bahwa seseorang yang rutin mengonsumsi obat antiepilepsi akan menurunkan risiko terjadinya kejang secara berulang.
6. Pemicu kejang merupakan salah satu faktor yang dapat memunculkan serangan kejang secara tiba-tiba pada penderita epilepsi, pada setiap penderita memiliki pemicu yang berbeda-beda tetapi ada juga yang tidak memiliki faktor pemicu.
7. Kualitas tidur, kualitas tidur yang buruk dapat meningkatkan frekuensi kejang 10 kali lebih besar (Nisa, 2015). Kualitas tidur yang baik dapat dilihat dari kontinuitas atau waktu tidur yang tidak berhenti hingga bangun esok hari.
8. Kondisi tempat tinggal, tekanan atmosfer yang rendah dan kelembaban udara relatif tinggi dapat meningkatkan risiko kejang epilepsi, sedangkan suhu lingkungan yang tinggi bisa mengurangi risiko kejang (Rakers dkk., 2017).

Agar lebih jelas mengenai gambaran teori dan konsep dari kejang epilepsi, maka disajikan kerangka konsep pada Gambar 3.2.



Gambar 3.2 Kerangka Konsep Penelitian

Keterangan:

* : Variabel yang digunakan dalam analisis

Catatan:

Variabel pada kerangka konsep yang tidak diteliti disebabkan karena adanya data yang tidak lengkap.

3.4 Instrumen Penelitian

Penelitian ini menggunakan instrumen metode angket atau kuisisioner. Metode ini dipilih karena memiliki keuntungan yaitu responden dapat bebas menjawab dengan jujur dan tidak ada rasa malu. Pertanyaan yang diberikan sesuai pada Tabel 3.2.

Tabel 3.2 Aspek dan Pertanyaan Penelitian

Aspek Penelitian	Pertanyaan
Identitas Responden	<ol style="list-style-type: none"> 1. Nama panggilan 2. Jenis kelamin dan usia 3. Agama 4. Hobi 5. Alamat rumah dan nomor <i>handphone</i>
Faktor Sosio-Ekonomi	<ol style="list-style-type: none"> 1. Status pernikahan, pendidikan terakhir 2. Pekerjaan dan pengeluaran bulanan
Faktor Riwayat Kesehatan	<ol style="list-style-type: none"> 1. Usia terdiagnosis epilepsi 2. Penyebab penyakit epilepsi 3. Riwayat keluarga, perinatal, demam disertai kejang, trauma kepala, terapi 4. Frekuensi kejang per hari, bulan, tahun 5. Pemicu serangan kejang 6. Kemampuan merasakan aura 7. Tindakan preventif yang dilakukan
Faktor Gaya Hidup	<ol style="list-style-type: none"> 1. Intensitas melakukan ibadah 2. Konsumsi alkohol, rokok 3. Konsumsi obat antiepilepsi 4. Keberanian mengemudi 5. Memerlukan pendampingan 6. Kebiasaan olah tubuh 7. Makanan atau minuman yang dihindari 8. Lama tidur dan intensitas terbangun 9. Pemakaian AC di rumah dan di kantor

Kuisisioner yang disusun dikonsultasikan dengan Komunitas ODE dan Dokter Poli Neurologi Rumah Sakit Universitas Airlangga. Kuisisioner tersebut bertujuan untuk mencari informasi mengenai

identitas, riwayat kesehatan, kondisi sosio-ekonomi, dan gaya hidup dari responden selama menderita epilepsi. Pertanyaan yang diberikan terdiri dari pertanyaan terbuka dan pertanyaan tertutup yang menggunakan skala nominal, dan skala ordinal dengan pilihan jawaban berdasarkan informasi dari hal-hal yang dialami oleh penderita epilepsi pada umumnya. Agar lebih jelas mengenai kuisioner penelitian yang disusun dapat dilihat pada Lampiran 1

3.5 Variabel Penelitian dan Definisi Operasional

Tidak semua pertanyaan pada instrumen memiliki pengaruh terhadap kejadian kejang penderita epilepsi, sehingga pada penelitian ini hanya digunakan beberapa pertanyaan untuk menjadi variabel dalam analisis. Sedangkan variabel lain akan digunakan untuk mengetahui karakteristik dari penderita epilepsi. Variabel yang digunakan dalam analisis disajikan pada Tabel 3.3.

Tabel 3.3 Variabel Penelitian

Variabel	Keterangan	Skala	Kategori
y	Kejadian kejang dalam waktu satu bulan	Nominal	1: Tidak mengalami kejang 2: Mengalami kejang
x_1	Jenis kelamin	Nominal	1: Laki-laki 2: Perempuan
x_2	Usia	Rasio	-
x_3	Usia terdiagnosis	Rasio	-
x_4	Riwayat demam kejang	Nominal	1: Tidak ada 2: Ada
x_5	Konsumsi obat anti epilepsi	Nominal	1: Tidak 2: Ya
x_6	Pemicu kejang	Nominal	1: Tidak ada 2: Ada
x_7	Intensitas terbangun	Ordinal	1: < 1 kali 2: \geq 1 kali
x_8	Pemakaian AC	Nominal	1: Tidak memakai AC 2: Memakai AC

Variabel yang digunakan dalam analisis terdiri dari satu variabel respon dan delapan variabel prediktor. Adapun definisi operasional untuk masing-masing variabel yang digunakan adalah sebagai berikut:

a. Variabel Respon

Penelitian ini menggunakan variabel respon kejadian kejang yang dialami penderita epilepsi. Kejang yang diukur adalah kejang yang datang secara tiba-tiba dimana penderita akan jatuh lalu diikuti dengan gerakan kejang, dan tidak sadarkan diri. Dalam penelitian ini kejadian kejang diukur selama satu bulan dengan dua kategori yaitu tidak mengalami kejang dan mengalami kejang.

1. Tidak mengalami kejang yaitu penderita epilepsi yang tidak mengalami kejang sama sekali dalam waktu satu bulan.
2. Mengalami kejang yaitu penderita epilepsi yang mengalami kejang minimal satu kali dalam waktu satu bulan

b. Variabel Prediktor

Penelitian ini menggunakan delapan variabel prediktor yang diduga berpengaruh terhadap kejadian kejang penderita epilepsi.

1. Jenis kelamin penderita epilepsi (x_1), dibagi menjadi dua kategori yaitu laki-laki dan perempuan.
2. Usia (x_2) yang digunakan adalah usia saat ulang tahun terakhir dengan perhitungan usia berdasarkan pada kalender Masehi, yang dilakukan pembulatan ke bawah.
3. Usia terdiagnosis (x_3) merupakan usia pertama kali penderita dinyatakan menderita penyakit epilepsi. Dimana untuk pencatatan dilakukan pembulatan ke bawah dengan satuan tahun.
4. Riwayat demam disertai kejang (x_4) merupakan ada atau tidaknya kejadian kenaikan suhu tubuh yang tinggi dan diikuti oleh kejang yang dialami oleh penderita epilepsi
5. Konsumsi obat anti epilepsi (x_5) dibedakan menjadi dua kategori yaitu tidak mengonsumsi obat dan mengonsumsi obat
6. Faktor pemicu kejang (x_6) merupakan ada atau tidaknya faktor yang bisa menyebabkan seseorang mengalami kejadian kejang
7. Intensitas terbangun (x_7) adalah berapa kali penderita terbangun atau terjaga diantara waktu tidur di malam hari, yang dibagi menjadi dua kategori yaitu kurang dari atau sama dengan 1 kali dan lebih dari satu kali terbangun

8. Pemakaian AC (x_8) pada tempat tinggal adalah ada atau tidaknya pemakaian AC di dalam rumah, dilakukan pengkategorian menjadi dua yaitu tidak memakai AC dan memakai AC.

3.6 Struktur Data Penelitian

Struktur data yang digunakan dalam penelitian ini ditunjukkan pada Tabel 3.4 merupakan struktur data dengan jumlah n penderita epilepsi. Dimana variabel respon (y) yang digunakan memiliki dua kategori yaitu tidak mengalami kejang dan mengalami kejang. Sedangkan variabel prediktor (x) yang digunakan ada sebanyak 8 variabel yang merupakan faktor-faktor risiko yang diduga mempengaruhi serangan kejang pada penderita epilepsi.

Tabel 3.4 Struktur Data Penelitian

Observasi	y	x_1	x_2	...	x_8
1	y_1	$x_{1,1}$	$x_{1,2}$...	$x_{1,8}$
2	y_2	$x_{2,1}$	$x_{2,2}$...	$x_{2,8}$
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
n	y_n	$x_{n,1}$	$x_{n,3}$...	$x_{n,8}$

3.7 Langkah Penelitian

Langkah-langkah penelitian akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Menyusun kuisisioner dengan pertanyaan dan jawaban berdasarkan studi literatur pada penelitian terdahulu dan hal-hal yang umumnya dialami oleh penderita epilepsi. Selanjutnya kuisisioner tersebut dikonsultasikan dengan Komunitas ODE dan Dokter Poli Neurologi Rumah Sakit Universitas Airlangga.
2. Melakukan pengambilan data terhadap penderita epilepsi yang terpilih sebagai sampel penelitian yaitu sebanyak 51 penderita epilepsi dari Komunitas ODE.
3. Melakukan *pre-processing* data untuk memilih variabel-variabel yang akan digunakan dalam analisis.
4. Melakukan analisis data berdasarkan informasi yang diperoleh dari 51 penderita epilepsi, dengan tahapan analisis sebagai berikut:

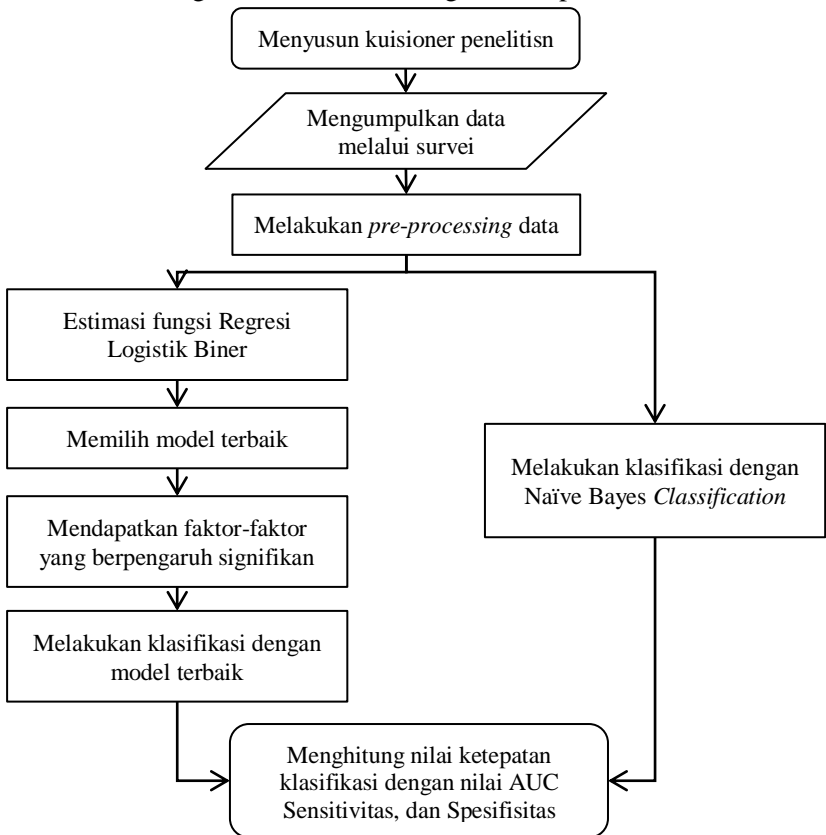
- a. Membuat deskripsi tentang karakteristik 51 penderita epilepsi pada Komunitas ODE dengan menggunakan analisis statistika deskriptif.
- b. Melakukan partisi data menjadi data training dan data testing menggunakan metode *5-fold cross validation*.
- c. Melakukan analisis regresi logistik biner untuk mendapatkan faktor-faktor yang mempengaruhi kejadian kejang dan mengklasifikasikan kejadian kejang pada 51 penderita epilepsi dengan tahapan analisis sebagai berikut:
 - Melakukan uji independensi untuk melihat ada atau tidaknya hubungan antar variabel prediktor (x) terhadap variabel respon (y) menggunakan persamaan (2.1).
 - Menghitung estimasi parameter regresi logistik biner.
 - Melakukan seleksi variabel untuk mendapatkan model terbaik dengan metode *backward elimination*.
 - Melakukan pengujian signifikansi parameter secara serentak menggunakan Uji *Ratio Likelihood* sesuai persamaan (2.12), dan secara parsial dengan Uji Wald menggunakan persamaan (2.13).
 - Melakukan pembentukan model dan menginterpretasikan model yang terbentuk.
 - Menghitung dan menginterpretasikan nilai *odds ratio* sesuai dengan persamaan (2.14).
 - Melakukan uji kesesuaian model menggunakan statistik Hosmer and Lemeshow *test* pada persamaan (2.15).
 - Menghitung ketepatan klasifikasi model menggunakan persamaan (2.20)-(2.22).
- d. Melakukan klasifikasi kejadian kejang 51 penderita epilepsi menggunakan naïve bayes dengan tahapan sebagai berikut:
 - Menghitung probabilitas awal (*prior probability*) untuk menentukan peluang dari setiap kategori kejadian kejang.
 - Melakukan perhitungan nilai probabilitas bersyarat untuk setiap variabel j , $j = 1, 2, \dots, p$ sesuai persamaan (2.17). Apabila ada data yang kontinyu, maka akan dihitung nilai rata-rata, varians, dan standar deviasi dari masing-masing variabel pada setiap kategori, dimana nilai-nilai tersebut

akan digunakan untuk menghitung probabilitas menggunakan pendekatan distribusi normal sesuai persamaan (2.18).

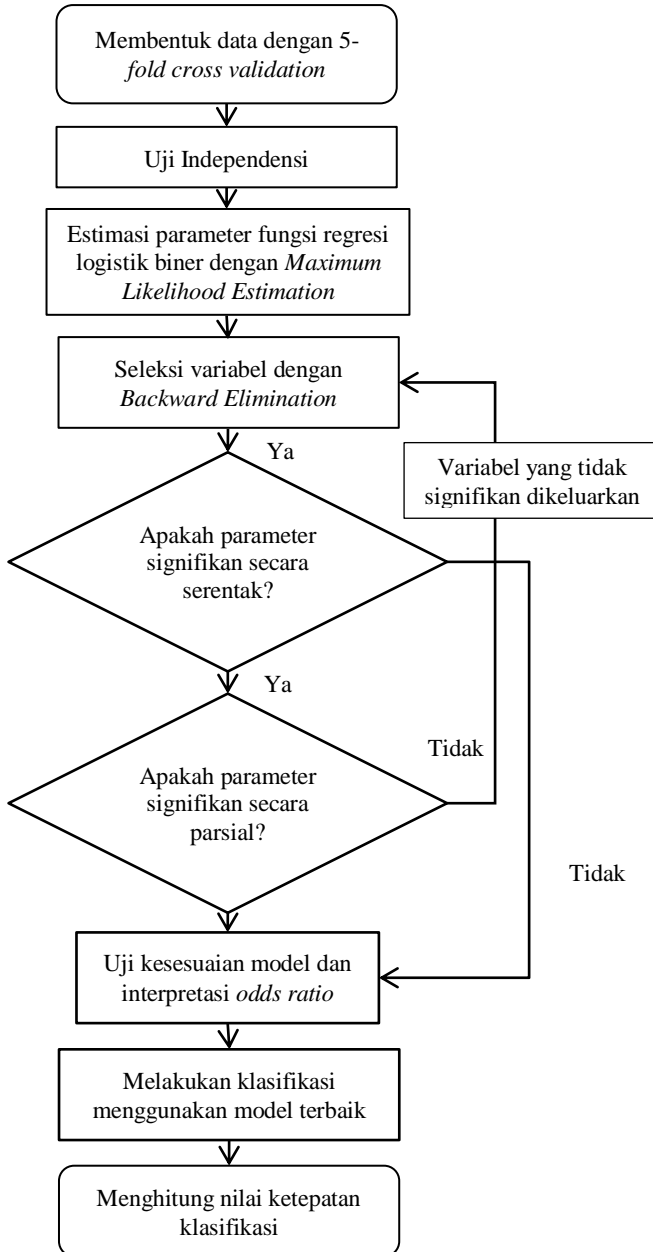
- Menentukan probabilitas akhir (*posterior probability*) dari semua parameter pada setiap kategori sesuai persamaan (2.16).
- Menentukan prediksi kelompok kategori berdasarkan nilai probabilitas tertinggi.
- Menghitung ketepatan klasifikasi menggunakan persamaan (2.20)-(2.23)

5. Menarik kesimpulan dari hasil analisis.

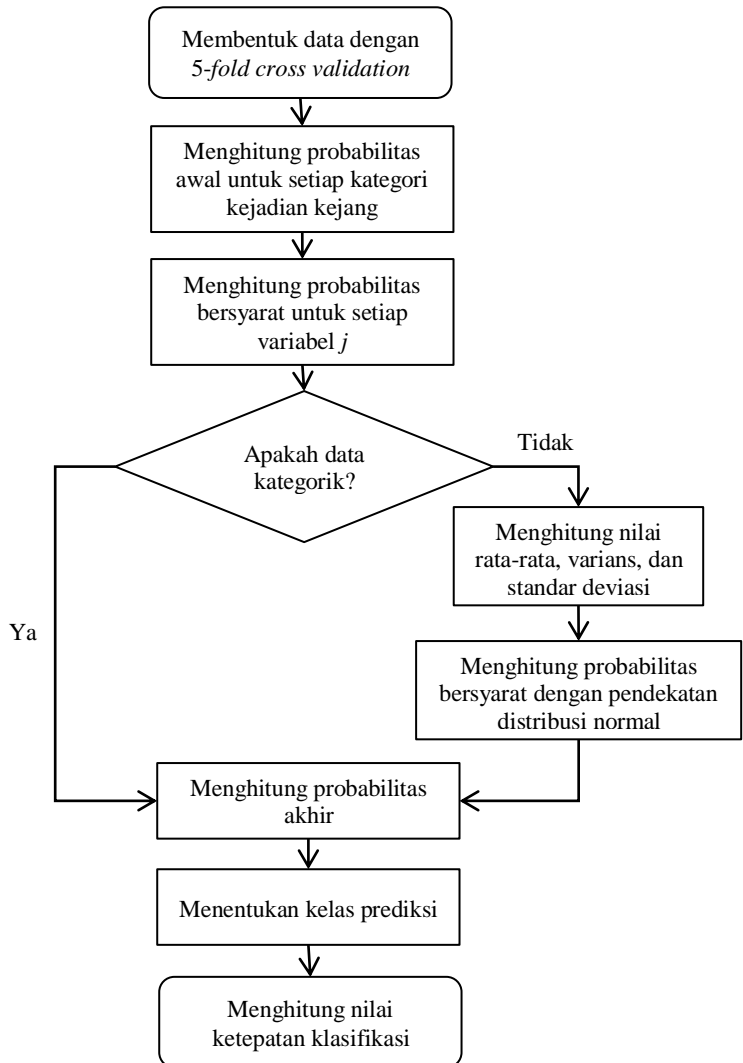
Langkah-langkah analisis yang akan dilakukan pada penelitian ini secara umum digambarkan dalam diagram alir pada Gambar 3.3-3.5.



Gambar 3.3 Diagram Alir Penelitian



Gambar 3.4 Diagram Alir Klasifikasi Menggunakan Regresi Logistik Biner



Gambar 3.4 Diagram Alir Klasifikasi Menggunakan Naïve Bayes

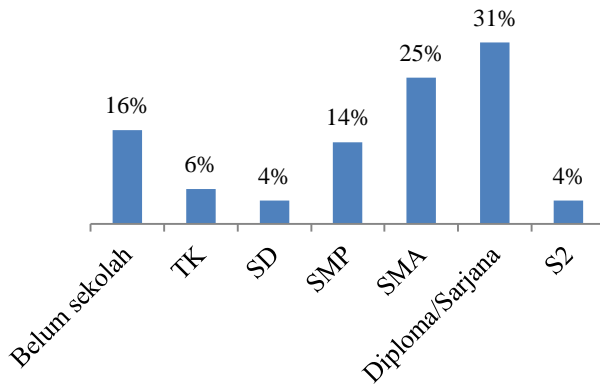
(Halaman ini sengaja dikosongkan)

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab ini akan dijelaskan mengenai hasil analisis dan pembahasan berdasarkan pengolahan data yang telah dilakukan. Metode yang digunakan dalam analisis ini meliputi statistika deskriptif untuk menggambarkan karakteristik penderita epilepsi serta analisis regresi logistik biner dan naïve bayes untuk memodelkan serta mengklasifikasikan kejadian kejang penderita epilepsi berdasarkan faktor-faktor risiko kejang epilepsi.

4.1 Karakteristik Penderita Epilepsi

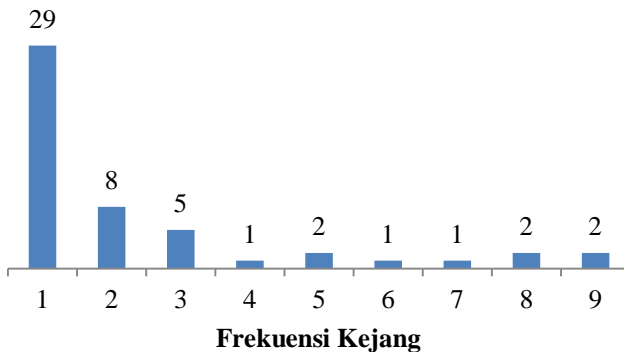
Analisis karakteristik dilakukan untuk mendeskripsikan pendidikan, kejadian kejang, dan riwayat kesehatan penderita epilepsi dalam Komunitas ODE. Sebagian besar penderita epilepsi dalam Komunitas ODE yang menjadi responden penelitian memiliki pendidikan terakhir sebagai seorang diploma/sarjana. Terdapat 31% dari 51 penderita epilepsi atau sebanyak 16 penderita epilepsi dalam Komunitas ODE yang memiliki pendidikan terakhir diploma/sarjana dan terdapat 4% atau 2 penderita epilepsi yang merupakan lulusan S2 seperti yang ditunjukkan pada Gambar 4.1. Sedangkan sisanya merupakan lulusan pada jenjang pendidikan lain seperti SMA, SMP, SD, dan TK, serta terdapat penderita yang belum sekolah.



Gambar 4.1 Proporsi Penderita Epilepsi Berdasarkan Pendidikan Terakhir

Seseorang yang memiliki latar belakang pendidikan tinggi akan lebih sadar terhadap gejala-gejala sebelum terjadinya kejang, sehingga dapat mengendalikan diri agar tidak sampai terjadi kejang dan juga dapat mengontrol banyaknya kejang yang dialaminya. Artinya pemahaman mengenai frekuensi kejang yang dialami sejalan dengan pendidikan yang dimiliki oleh penderita epilepsi.

Selanjutnya pada Gambar 4.2 menunjukkan jumlah penderita epilepsi berdasarkan frekuensi kejang yang dialami satu bulan terakhir.

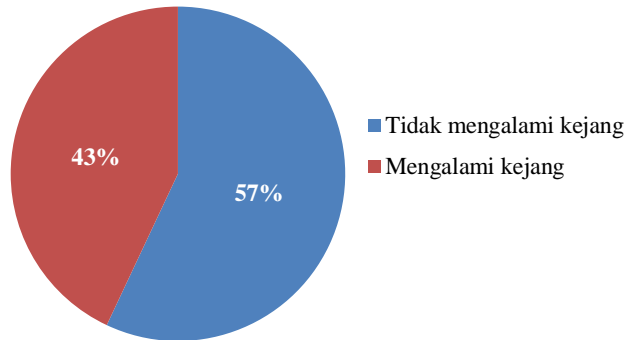


Gambar 4.2 Jumlah Penderita Epilepsi Berdasarkan Frekuensi Kejang

Gambar 4.2 menunjukkan banyaknya penderita epilepsi dari masing-masing frekuensi kejang yang dialami. Sebagian besar atau sebanyak 29 dari 51 penderita epilepsi dalam Komunitas ODE tidak mengalami kejang dalam kurun waktu satu bulan terakhir. Sedangkan sisanya sebanyak 22 penderita epilepsi mengalami kejang dengan frekuensi yang berbeda-beda. Frekuensi kejang terbanyak yaitu sebanyak 20 kali kejang dalam satu bulan dialami oleh 2 penderita epilepsi yang masih dalam kategori balita. Gambaran distribusi data frekuensi kejang tersebut menjadi salah satu pertimbangan untuk pembagian kelompok kejadian kejang. Kejadian kejang akan dikelompokkan menjadi dua kategori, yaitu tidak mengalami kejang dan mengalami kejang. Tidak mengalami kejang yaitu kelompok untuk penderita epilepsi yang tidak mengalami kejang sama sekali dalam waktu satu bulan terakhir, sedangkan

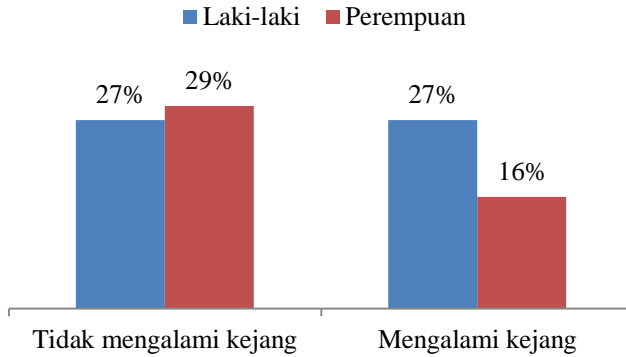
mengalami kejang merupakan kelompok bagi penderita epilepsi yang mengalami kejang satu kali atau lebih.

Gambar 4.3 memberikan gambaran mengenai proporsi jumlah penderita epilepsi yang terdapat pada masing-masing kategori kejadian kejang yaitu tidak mengalami kejang dan mengalami kejang. Antara jumlah penderita epilepsi yang mengalami maupun yang tidak mengalami kejang dalam waktu satu bulan terakhir memiliki proporsi yang hampir sama. Proporsi penderita epilepsi yang tidak mengalami kejang dalam jangka waktu satu bulan ada sebanyak 57%, sedangkan proporsi penderita epilepsi yang mengalami kejang dalam waktu satu bulan ada sebanyak 43%.



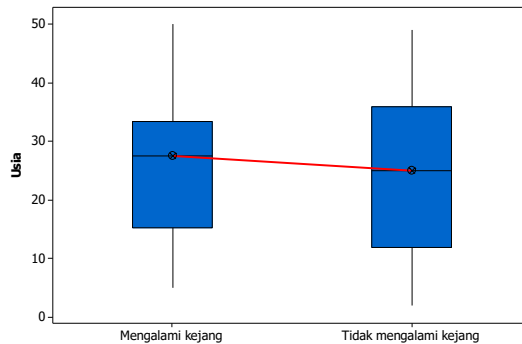
Gambar 4.3 Proporsi Penderita Epilepsi pada Tiap Kategori Kejadian Kejang

Proporsi jumlah penderita epilepsi laki-laki dan perempuan pada setiap kategori kejadian kejang ditunjukkan dengan diagram batang yang terdapat pada Gambar 4.4. Pada Gambar 4.4 ditunjukkan bahwa secara keseluruhan jumlah penderita epilepsi dengan jenis kelamin laki-laki lebih banyak dibandingkan jumlah penderita epilepsi yang berjenis kelamin perempuan. Pada kategori tidak mengalami kejang proporsi laki-laki dengan perempuan berbeda selisih 2% dengan proporsi lebih banyak perempuan yaitu 29% dari jumlah seluruh penderita epilepsi. Berbeda halnya dengan kategori mengalami kejang, terdapat perbedaan jumlah penderita epilepsi laki-laki yang hampir 2 kali lipat dari jumlah penderita epilepsi perempuan.



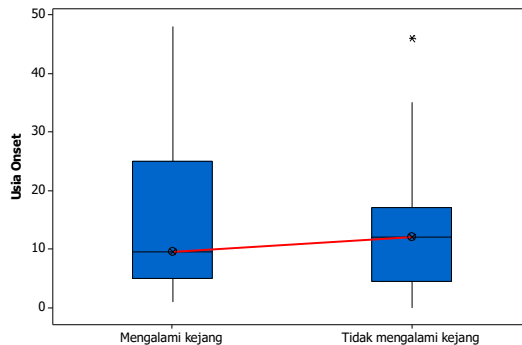
Gambar 4.4 Proporsi Penderita Epilepsi Berdasarkan Jenis Kelamin Antar Kategori Kejadian Kejang

Nilai tengah dari suatu variabel independen dapat digunakan untuk mengetahui karakteristik tertentu pada masing-masing kategori kejadian kejang. Ada terdapat beberapa ukuran nilai tengah, diantaranya adalah median. Perbedaan median dari variabel independen antar kategori kejadian kejang dapat dilihat menggunakan *boxplot*. *Boxplot* tersebut dapat menunjukkan adanya perbedaan dari setiap kategori kejadian kejang yang dialami oleh penderita epilepsi untuk variabel independen yang bersifat kontinyu. Pada penelitian ini variabel independen yang bersifat kontinyu ada dua yaitu variabel usia penderita epilepsi dan usia terdiagnosis penyakit epilepsi. *Boxplot* dari kedua variabel tersebut disajikan pada Gambar 4.5 dan Gambar 4.6. Pada Gambar 4.5 terlihat bahwa usia penderita epilepsi saat ini berbeda pada masing-masing kategori. Hal tersebut dapat ditunjukkan oleh garis median pada *boxplot* yang tidak sejajar. Pada pada kategori tidak mengalami kejang memiliki distribusi yang simetris, sedangkan kategori mengalami kejang menunjukkan distribusi yang tidak simetris. Hal ini dikarenakan penderita epilepsi dalam kategori mengalami kejang didominasi oleh penderita epilepsi yang berada pada usia remaja dan dewasa awal daripada penderita epilepsi usia anak.



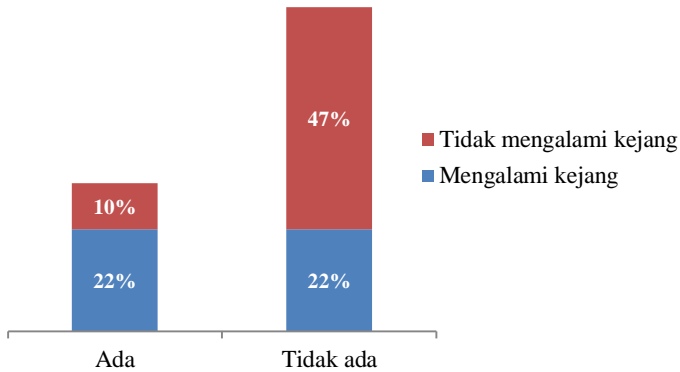
Gambar 4.5 Perbedaan Median Usia Penderita Epilepsi Antar Kategori Kejadian Kejang

Sedangkan Gambar 4.6 menunjukkan *boxplot* untuk usia terdiagnosis epilepsi pada setiap kategori kejadian kejang. Berdasarkan *boxplot* tersebut terlihat bahwa median usia terdiagnosis epilepsi pada masing-masing kategori kejadian kejang berbeda. Hal tersebut dapat dilihat dari garis median pada kedua *boxplot* yang tidak sejajar. Berbeda halnya dengan variabel usia penderita epilepsi saat ini, pada variabel usia terdiagnosis untuk kategori mengalami kejadian kejang didominasi dengan usia terdiagnosis yang lebih muda dikarenakan garis median yang tidak berada di tengah-tengah dan lebih condong berada di bawah.



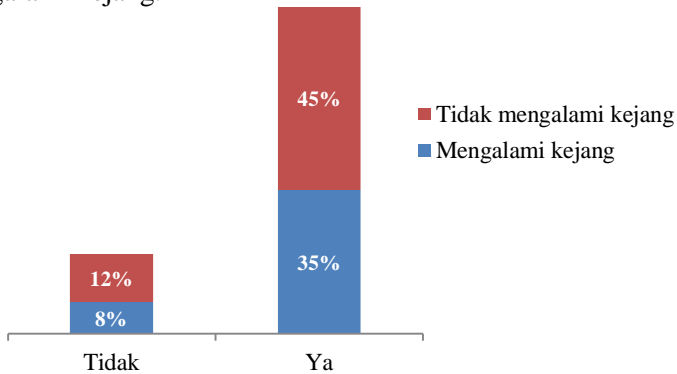
Gambar 4.6 Perbedaan Median Usia Terdiagnosis Penderita Epilepsi Antar Kategori Kejadian Kejang

Stacked bar chart dapat digunakan untuk melihat perbandingan data sejenis dalam satu item. Gambar 4.7 merupakan *stacked column* untuk variabel riwayat demam kejang.



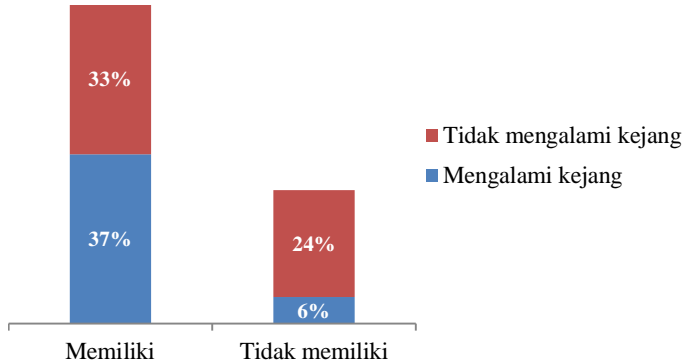
Gambar 4.7 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Riwayat Demam Disertai Kejang

Pada Gambar 4.7 terlihat bahwa penderita yang terpilih menjadi responden lebih banyak yang tidak ada riwayat demam kejang. Pada kategori ada riwayat demam kejang, proporsi penderita epilepsi yang mengalami kejang lebih banyak daripada penderita yang tidak mengalami kejang. Berbeda halnya pada kategori tidak ada riwayat demam kejang, dimana proporsi penderita yang tidak mengalami kejang dalam waktu satu bulan terakhir lebih banyak daripada yang mengalami kejang.



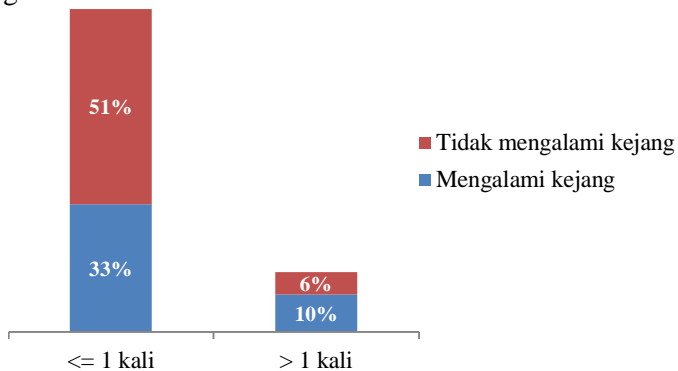
Gambar 4.8 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Kepatuhan Konsumsi Obat Antiepilepsi

Berdasarkan Gambar 4.8 diketahui bahwa ada penderita epilepsi yang tidak mengonsumsi obat antiepilepsi memiliki proporsi yang jauh lebih sedikit jika dibandingkan dengan penderita epilepsi yang mengonsumsi obat. Pada kategori penderita yang mengonsumsi obat, kejadian tidak mengalami kejang lebih besar proporsinya jika dibandingkan dengan yang mengalami kejang.



Gambar 4.9 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Pemicu Kejang

Berdasarkan Gambar 4.9 diketahui bahwa ada penderita epilepsi dalam Komunitas ODE yang memiliki faktor pemicu lebih banyak daripada yang tidak memiliki faktor pemicu kejang. Pada kategori penderita yang tidak memiliki pemicu kejang, proporsi penderita tidak mengalami kejang lebih banyak daripada yang mengalami kejang.



Gambar 4.10 Perbedaan Kejadian Kejang Penderita Epilepsi Berdasarkan Intensitas Terbangun

Berdasarkan Gambar 4.10 diketahui bahwa pada kategori intensitas terbangun saat malam hari kurang dari sama dengan satu kali, memiliki proporsi penderita yang tidak mengalami kejang lebih banyak daripada yang mengalami kejang. Berbeda dengan kategori intensitas terbangun lebih dari satu kali, dimana proporsi penderita epilepsi yang mengalami kejang lebih banyak.

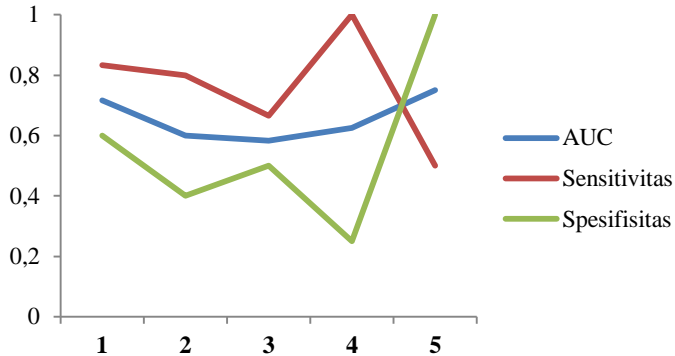
4.2 Analisis Regresi Logistik Biner Berdasarkan Faktor-Faktor yang Diduga Mempengaruhi Kejadian Kejang

Analisis regresi logistik biner bertujuan untuk mengklasifikasikan kejadian kejang pada 51 penderita epilepsi, dimana pada tahapan klasifikasinya akan diperoleh model dan variabel yang signifikan berpengaruh dalam mengklasifikasikan kejadian kejang pada penderita epilepsi. Sebelum melakukan tahapan analisis regresi logistik biner, terlebih dahulu dilakukan partisi data menggunakan *5-fold cross validation*. Teknik partisi ini membagi data menjadi dua bagian yaitu data *training* dan data *testing*. Data *training* digunakan untuk membuat model klasifikasi, sedangkan data *testing* digunakan untuk memeriksa ketepatan dari model yang terbentuk. Ketepatan model klasifikasi dapat diukur menggunakan nilai AUC (*Area Under Curve ROC*), sensitivitas, dan spesifisitas. Tabel 4.1 menyajikan nilai AUC, sensitivitas, dan spesifisitas dari data *testing* untuk masing-masing *fold*.

Tabel 4.1 Nilai AUC, Sensitivitas, dan Spesifisitas dari Data *Testing* Tiap *Fold*

<i>Fold</i> ke-	AUC	Sensitivitas	Spesifisitas
1	0,7167	0,8333	0,60
2	0,6000	0,8000	0,40
3	0,5833	0,6667	0,50
4	0,6250	1,0000	0,25
5	0,7500	0,5000	1,00

Agar lebih mudah untuk membandingkan hasil ketepatan klasifikasi yang didapatkan maka disajikan grafik pembandingan untuk nilai AUC, sensitivitas, dan spesifisitas dari masing-masing *fold* pada Gambar 4.11.



Gambar 4.11 Perbandingan Ketepatan Klasifikasi Tiap *Fold* pada Data *Testing* dengan Metode Regresi Logistik Biner

Gambar 4.11 menunjukkan perbedaan nilai AUC, sensitivitas, dan spesifisitas untuk masing-masing *fold*. Pada gambar tersebut terlihat jelas bahwa terdapat selisih ketepatan klasifikasi pada beberapa *fold*. Namun, nilai ketepatan klasifikasi pada *fold* kelima yang dilihat berdasarkan kriteria nilai AUC memiliki nilai yang lebih tinggi daripada *fold* lainnya. Sehingga untuk tahapan analisis regresi logistik digunakan *fold* yang memiliki nilai AUC tertinggi yaitu data *training* dan *testing* pada *fold* ke-5.

4.2.1 Uji Independensi Antara Kejadian Kejang dengan Faktor-Faktor Risiko Kejadian Kejang

Uji independensi bertujuan untuk mendeteksi hubungan antara kejadian kejang dengan faktor-faktor yang diduga berpengaruh terhadap kejadian kejang dengan hasil sesuai pada Tabel 4.2.

Tabel 4.2 Hasil Uji Independensi Antara Kejadian Kejang dengan Faktor Risiko

Variabel	χ^2_{hitung}	<i>p-value</i>
Jenis kelamin	0,004	0,951
Riwayat demam kejang	3,570	0,059*
Konsumsi obat	0,082	0,775
Pemicu kejang	5,487	0,019*
Intensitas terbangun	0,601	0,438
Pemakaian AC	0,854	0,355

*Sign. pada $\alpha = 0,1$

Hipotesis yang digunakan pada pengujian independensi tercantum pada sub bab 2.2. Hasil pada Tabel 4.2 menunjukkan bahwa terdapat dua variabel yang menghasilkan keputusan tolak H_0 karena variabel tersebut memiliki p -value yang kurang dari $\alpha = 0,1$, sehingga dapat diambil kesimpulan bahwa variabel yang memiliki hubungan dengan kejadian kejang pada penderita epilepsi adalah variabel riwayat demam disertai kejang dan variabel pemicu kejang.

4.2.2 Pembentukan Model Regresi Logistik Biner dengan Seluruh Variabel

Setelah dilakukan pengujian independensi untuk mengidentifikasi hubungan antara faktor-faktor risiko terhadap kejadian kejang penderita epilepsi, selanjutnya dilakukan pembentukan model untuk seluruh variabel prediktor terhadap variabel respon. Berikut adalah model logit yang didapatkan dari analisis regresi logistik biner.

$$\hat{g}(x) = -1,961 + 0,520 \text{ Jenis Kelamin}_{(2)} - 0,114 \text{ Usia} + 0,120 \text{ Usia Terdiagnosis} + 2,169 \text{ Riwayat Demam Kejang}_{(2)} - 0,167 \text{ Konsumsi Obat}_{(2)} + 2,670 \text{ Pemicu Kejang}_{(2)} + 0,251 \text{ Intensitas Terbangun}_{(2)} + 0,574 \text{ Pemakaian AC}_{(2)}$$

Pengujian signifikansi parameter merupakan tahapan selanjutnya yang harus dilakukan untuk mengetahui variabel mana saja yang berpengaruh signifikan secara serentak dan parsial terhadap variabel kejadian kejang. Berikut adalah hasil signifikansi parameter variabel kejadian kejang dengan faktor-faktor yang diduga mempengaruhi.

a. Uji Serentak terhadap Faktor-Faktor Risiko Kejadian Kejang Penderita Epilepsi

Uji serentak dilakukan untuk mengetahui pengaruh variabel prediktor terhadap kejadian kejang penderita epilepsi secara serentak dengan memasukkan semua variabel. Hipotesis yang digunakan adalah yang tercantum pada sub bab 2.5 (bagian a). Tabel 4.3 adalah asil uji serentak dengan memasukkan seluruh variabel prediktor.

Tabel 4.3 Hasil Uji Serentak Regresi Logistik Biner Tahap Awal

χ^2_{hitung}	db	p -value
15,226	8	0,055*

*Sign. pada $\alpha = 0,1$

Tabel 4.3 menunjukkan bahwa keputusan uji serentak pada kejadian kejang penderita epilepsi dengan faktor-faktor yang diduga mempengaruhinya adalah tolak H_0 karena memiliki $p\text{-value} < \alpha$ atau $0,055 < 0,1$, sehingga dapat disimpulkan bahwa minimal terdapat satu dari variabel prediktor yang berpengaruh signifikan terhadap kejadian kejang pada penderita epilepsi. Pengujian dilanjutkan dengan uji parsial untuk mendeteksi variabel yang berpengaruh terhadap kejadian kejang penderita epilepsi.

b. Uji Parsial terhadap Faktor-Faktor Risiko Kejadian Kejang Penderita Epilepsi

Pengujian secara parsial dilakukan untuk mengetahui variabel prediktor mana yang berpengaruh signifikan dalam model pada taraf signifikansi 0,1 berdasarkan $p\text{-value}$ yang disajikan pada Tabel 4.3 dengan menggunakan hipotesis yang sesuai pada sub bab 2.5 (bagian b).

Tabel 4.4 Hasil Uji Parsial Regresi Logistik Biner Tahap Awal

Variabel	B	Wald	$p\text{-value}$	90% CI	
				Lower	Upper
Jenis kelamin ₍₂₎	0,520	0,364	0,546	0,407	6,954
Usia	-0,114	3,791	0,052*	0,810	0,982
Usia terdiagnosis	0,120	3,344	0,067*	1,012	1,256
Riwayat demam kejang ₍₂₎	2,169	3,885	0,049*	1,432	53,494
Konsumsi obat ₍₂₎	-0,167	0,033	0,855	0,187	3,823
Pemicu kejang ₍₂₎	2,670	6,607	0,010*	2,616	79,754
Intensitas terbangun ₍₂₎	0,251	0,050	0,823	0,203	8,151
Pemakaian AC ₍₂₎	0,574	0,425	0,514	0,417	7,555

*Sign. pada $\alpha = 0,1$

Berdasarkan hasil uji parsial pada Tabel 4.4, diketahui bahwa terdapat empat variabel prediktor yang memiliki $p\text{-value}$ kurang dari $\alpha = 0,1$ yaitu variabel usia, usia terdiagnosis, riwayat demam kejang dan pemicu kejang dengan $p\text{-value}$ masing-masing yaitu 0,052; 0,067; 0,049; dan 0,010. Sehingga dapat disimpulkan bahwa variabel usia, usia terdiagnosis, riwayat demam kejang dan

pemicu kejang memiliki pengaruh yang signifikan terhadap kejadian kejang yang dialami penderita epilepsi, sedangkan keempat variabel lain yaitu variabel jenis kelamin, konsumsi obat, intensitas terbangun, dan pemakaian AC tidak memiliki pengaruh yang signifikan terhadap kejadian kejang epilepsi. Selanjutnya empat variabel yang tidak memiliki pengaruh yang signifikan terhadap kejadian kejang pada penderita epilepsi akan dikeluarkan satu per satu dari model sesuai dengan tahapan metode *backward elimination*.

4.2.3 Pembentukan Model Regresi Logistik Biner dengan *Backward Elimination*

Metode *backward* digunakan untuk seleksi variabel dengan cara memasukkan semua variabel prediktor ke dalam model terlebih dahulu. Seleksi variabel dilakukan berdasarkan nilai AIC, dimana model terbaik yang terpilih dari *fold* ke-5 adalah model yang memiliki AIC terkecil. Tahapan seleksi variabel menggunakan metode *backward* ditampilkan pada Lampiran 7. Variabel yang terpilih menggunakan metode *backward elimination* adalah usia (x_2), usia terdiagnosis (x_3), riwayat demam kejang (x_4) dan pemicu kejang (x_6). Berdasarkan variabel yang signifikan, dilakukan kembali pemodelan menggunakan regresi logistik biner dengan hasil sebagai berikut:

$$\hat{g}(x) = -1,518 - 0,110 \text{ Usia} + 0,117 \text{ Usia Terdiagnosis} + 1,939 \text{ Riwayat Demam Kejang}_{(2)} + 2,528 \text{ Pemicu Kejang}_{(2)}$$

a. Uji Serentak terhadap Faktor-Faktor yang Berpengaruh Signifikan pada Kejadian Kejang Penderita Epilepsi

Selanjutnya akan dilakukan pengujian signifikansi parameter kembali terhadap model yang didapatkan. Pengujian signifikansi parameter akan dilakukan secara serentak dan parsial. Hipotesis uji serentak yang digunakan sesuai pada sub bab 2.5 (bagian a). Sesuai dengan Lampiran 8 berikut adalah hasil uji serentak setelah dilakukan *backward elimination*.

Tabel 4.5 Hasil Uji Serentak Setelah Dilakukan *Backward Elimination*

χ^2_{hitung}	db	<i>p-value</i>
14,067	4	0,007*

*Sign. pada $\alpha = 0,1$

Tabel 4.5 menunjukkan bahwa pada pengujian serentak didapatkan nilai χ^2_{hitung} sebesar 14,067 dan p -value sebesar 0,007 dimana p -value tersebut signifikan pada $\alpha = 0,1$ maka dapat diputuskan bahwa minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap kejadian kejang pada penderita epilepsi.

b. Uji Parsial terhadap Faktor-Faktor yang Berpengaruh Signifikan pada Kejadian Kejang Penderita Epilepsi

Dikarenakan pada pengujian secara serentak diperoleh kesimpulan bahwa minimal terdapat satu variabel prediktor yang memiliki pengaruh yang signifikan terhadap variabel kejadian kejang, maka pengujian dilanjutkan dengan uji parsial untuk mengetahui ada atau tidaknya pengaruh yang signifikan secara parsial dari keempat variabel yang digunakan dalam pembentukan model terhadap kejadian kejang. Hipotesis yang digunakan tercantum dalam sub bab 2.5 (bagian b).

Tabel 4.6 Uji Parsial Setelah Dilakukan *Backward Elimination*

Variabel	B	Wald	p -value	90% CI		Exp(B)
				Lower	Upper	
Usia	-0,110	4,182	0,041*	0,820	0,979	0,896
Usia terdiagnosis	0,117	3,645	0,056*	1,016	1,244	1,124
Riwayat demam kejang ₍₂₎	1,930	3,808	0,051*	1,354	35,032	6,888
Pemicu kejang ₍₂₎	2,528	6,862	0,009*	2,561	61,238	12,524

*Sign. pada $\alpha = 0,1$

Tabel 4.6 menunjukkan bahwa variabel usia, usia terdiagnosis, riwayat demam disertai kejang, dan pemicu kejang berpengaruh nyata pada taraf signifikansi 0,1. Sehingga kejadian kejang pada penderita epilepsi ditentukan oleh kondisi variabel-variabel tersebut.

Hasil uji kesesuaian model menghasilkan nilai p -value sebesar 0,673, dengan menggunakan hipotesis sesuai pada sub bab 2.7 maka didapatkan kesimpulan bahwa model akhir yang terbentuk hasil dari *backward elimination* telah sesuai artinya tidak ada perbedaan yang nyata antara hasil observasi dengan kemungkinan hasil prediksi model. Selanjutnya interpretasi model dilakukan dengan

menggunakan nilai *odds ratio*. Nilai *odds ratio* yang didapatkan berasal dari variabel yang masuk dalam model logit terbaik yaitu variabel usia, usia terdiagnosis, riwayat demam kejang dan pemicu kejang, dimana nilai tersebut diperoleh dari hasil $\exp(B)$ sesuai pada Tabel 4.6. Berikut adalah interpretasi nilai *odds ratio* dari masing-masing variabel dengan menggunakan *first reference* untuk setiap kategorinya.

1. *Odds ratio* usia (x_2): $\Psi = e^{-0,110} = \mathbf{0,896}$, untuk memudahkan interpretasi maka nilai akan diubah menjadi $\mathbf{1/0,896 = 1,116}$. Nilai tersebut memiliki makna bahwa semakin mudanya usia yang dimiliki penderita epilepsi maka risiko untuk mengalami kejadian kejang dalam waktu satu bulan terakhir akan naik sebesar 1,116 kali. Dikarenakan nilai kenaikan risiko yang cukup kecil maka dapat dikatakan pula bahwa penderita epilepsi yang semakin muda memiliki risiko untuk mengalami kejang dalam waktu satu bulan terakhir adalah cenderung tetap.
2. *Odds ratio* usia terdiagnosis (x_3): $\Psi = e^{0,117} = \mathbf{1,124}$. Nilai *odds ratio* tersebut berarti bahwa jika seseorang terdiagnosis menderita penyakit epilepsi pada usia yang semakin tua maka risiko untuk mengalami kejadian kejang dalam waktu satu bulan terakhir akan meningkat sebesar 1,124 kali.
3. *Odds ratio* riwayat demam disertai kejang (x_4): $\Psi = e^{1,930} = \mathbf{6,888}$. Hal ini dapat diartikan bahwa penderita epilepsi yang memiliki riwayat demam disertai kejang memiliki risiko 6,9 kali lebih besar untuk mengalami kejadian kejang dalam waktu satu bulan terakhir daripada penderita epilepsi yang tidak memiliki riwayat demam disertai kejang.
4. *Odds ratio* pemicu kejang (x_6): $\Psi = e^{2,528} = \mathbf{12,524}$, berarti penderita epilepsi yang memiliki faktor pemicu kejang lebih berisiko untuk mengalami kejadian kejang dalam waktu satu bulan terakhir 12,5 kali lebih besar daripada penderita yang tidak memiliki faktor pemicu kejang.

Setelah didapatkan model terbaik, dapat dilakukan perhitungan nilai peluang untuk setiap variabel respon. Berikut adalah simulasi perhitungan nilai peluang untuk data *testing* pertama pada *fold* ke-5, penderita epilepsi berusia 17 tahun, terdiagnosis menderita epilepsi

usia 9 tahun, memiliki riwayat demam disertai kejang dan memiliki faktor pemicu kejang.

$$\begin{aligned}
 \pi_2 &= \frac{\exp[g(x)]}{1 + \exp[g(x)]} \\
 &= \frac{\exp[-1,518 - 0,110x_2 + 0,117x_3 + 1,930x_{4(2)} + 2,528x_{6(2)}]}{1 + \exp[-1,518 - 0,110x_2 + 0,117x_3 + 1,930x_{4(2)} + 2,528x_{6(2)}]} \\
 &= \frac{\exp[-1,518 - 0,110(17) + 0,117(9) + 1,930(1) + 2,528(1)]}{1 + \exp[-1,518 - 0,110(17) + 0,117(9) + 1,930(1) + 2,528(1)]} \\
 &= \frac{\exp[2,123]}{1 + \exp[2,123]} \\
 &= 0,8931 \\
 \pi_1 &= 1 - 0,8931 \\
 &= 0,1069
 \end{aligned}$$

Penderita epilepsi dengan kategori tersebut memiliki peluang untuk mengalami kejang adalah sebesar 0,8931, sedangkan peluang untuk tidak mengalami kejang adalah sebesar 0,1069, sehingga penderita tersebut terklasifikasikan pada kategori mengalami kejadian kejang. Perhitungan tersebut dilakukan untuk seluruh data *testing*, sehingga diperoleh hasil klasifikasi yang disajikan pada Tabel 4.7

Tabel 4.7 *Confusion Matrix* Regresi Logistik Biner Hasil Seleksi *Backward*

Observasi	Prediksi	
	Mengalami Kejang	Tidak Mengalami Kejang
Mengalami Kejang	3	3
Tidak Mengalami Kejang	0	4

Tabel 4.7 menunjukkan hasil klasifikasi kejadian kejang penderita pada data *testing*. Selanjutnya dilakukan perhitungan untuk mendapatkan nilai ketepatan klasifikasi berdasarkan nilai AUC, sensitifitas, dan spesifisitas seperti berikut:

$$\text{AUC} = \frac{1}{2} \left(\frac{3}{3+3} + \frac{4}{0+4} \right) = 0,75$$

$$\text{Sensitivitas} = \frac{3}{3+3} = 0,5$$

$$\text{Spesifisitas} = \frac{4}{0+4} = 1$$

Berdasarkan perhitungan tersebut dapat diketahui bahwa nilai AUC dari hasil pengklasifikasian adalah sebesar 75%, dimana hasil klasifikasi tersebut termasuk dalam kategori *fair classification*. Nilai sensitivitas yang dihasilkan atau penderita epilepsi yang mengalami kejadian kejang dan tepat diklasifikasikan ke dalam kategori mengalami kejadian kejang adalah sebesar 50%. Sedangkan nilai spesifisitas yang dihasilkan adalah sebesar 100%, nilai tersebut menunjukkan bahwa seluruh penderita epilepsi yang tidak mengalami kejang dalam waktu satu bulan terakhir diklasifikasikan secara tepat tidak mengalami kejadian kejang.

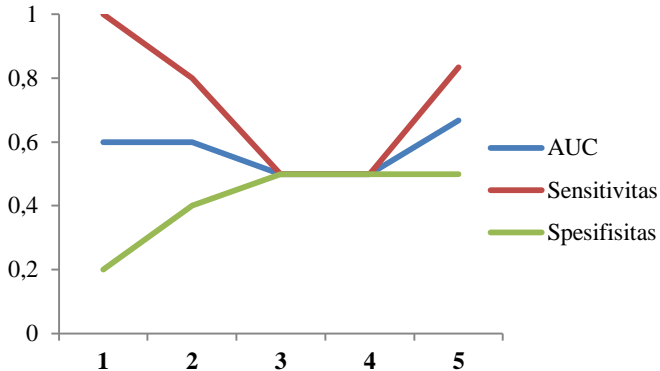
4.3 Klasifikasi Kejadian Kejang Penderita Epilepsi dengan Naïve Bayes

Langkah pertama yang dilakukan pada klasifikasi dengan naïve bayes adalah membagi data menjadi data *training* dan *testing* menggunakan *5-fold cross validation*. Kriteria yang digunakan untuk menilai ketepatan model klasifikasi adalah menggunakan nilai AUC, sensitivitas, dan spesifisitas. Perhitungan ketiga nilai tersebut sesuai dengan persamaan (2.20)-(2.22) yang terdapat pada sub bab 2.5. Nilai AUC, sensitivitas, spesifisitas dari data *testing* untuk masing-masing *fold* ditampilkan pada Tabel 4.8.

Tabel 4.8 Nilai AUC, Sensitivitas, dan Spesifisitas Tiap *Fold* pada Data *Testing*

<i>Fold</i> ke-	AUC	Sensitivitas	Spesifisitas
1	0,6000	1,0000	0,20
2	0,6000	0,8000	0,40
3	0,5000	0,5000	0,50
4	0,5000	0,5000	0,50
5	0,6667	0,8333	0,50

Agar lebih lebih jelas mengenai perbandingan ketepatan klasifikasi dari data testing untuk setiap *fold*, selanjutnya dibuat grafik perbandingan nilai AUC, sensitivitas, dan spesifisitas dari *fold* pertama sampai *fold* kelima sesuai pada Gambar 4.12.



Gambar 4.12 Perbandingan Ketepatan Klasifikasi Tiap *Fold* pada Data *Testing* dengan Metode Naïve Bayes

Gambar 4.12 menunjukkan bahwa hasil ketepatan klasifikasi dengan menggunakan metode naïve bayes yang diukur dengan menggunakan nilai AUC, sensitivitas, dan spesifisitas memiliki variasi hasil ketepatan klasifikasi dengan selisih yang tidak terlalu besar antar *fold* pengamatan. Berdasarkan grafik tersebut dapat diketahui pula bahwa *fold* yang menghasilkan nilai AUC dan spesifisitas tertinggi adalah pada *fold* kelima, namun untuk nilai sensitivitas tertinggi berada pada *fold* pertama. Dikarenakan kedua kriteria ketepatan klasifikasi (AUC dan spesifisitas) menunjukkan hasil yang tinggi pada *fold* kelima, maka dapat disimpulkan bahwa *fold* yang menghasilkan ketepatan klasifikasi tertinggi adalah *fold* kelima.

Data yang digunakan pada klasifikasi berikut merupakan data *training* dan *testing* pada *fold* ke-5 karena menghasilkan nilai akurasi tertinggi berdasarkan nilai AUC dan spesifisitas dimana model klasifikasinya melibatkan seluruh variabel. Langkah awal untuk melakukan klasifikasi dengan menggunakan naïve bayes adalah dengan menghitung nilai probabilitas awal. Data *training* yang digunakan sebanyak 41 observasi yang terdiri atas 23 observasi pada kategori tidak mengalami kejang dan 18 observasi pada kategori

mengalami kejang. Sehingga didapatkan nilai probabilitas awal seperti yang disajikan pada Tabel 4.9 yang didapatkan dengan membagi jumlah observasi pada masing-masing kategori kejang dengan keseluruhan observasi.

Tabel 4.9 Nilai Probabilitas Awal pada Data *Training*

Kategori Kejadian Kejang	Probabilitas
Tidak mengalami kejang	0,5609
Mengalami kejang	0,4391

Berdasarkan Tabel 4.9 dapat diketahui bahwa probabilitas penderita epilepsi yang masuk dalam kategori tidak mengalami kejang adalah sebesar 0,5609 dan probabilitas penderita epilepsi yang mengalami kejang sebesar 0,4391. Hasil probabilitas tersebut menunjukkan bahwa penderita epilepsi yang tidak mengalami kejang dan mengalami kejang dalam waktu satu bulan terakhir memiliki probabilitas yang hampir sama.

Langkah selanjutnya adalah menghitung nilai probabilitas bersyarat tiap kategori kejadian kejang untuk setiap variabel independen pada data *training*. Masing-masing nilai probabilitas bersyarat diperoleh dari pembagian jumlah observasi pada suatu kategori terhadap jumlah observasi keseluruhan dalam kategori tersebut. Nilai probabilitas bersyarat pada setiap variabel independen yang bersifat kategorik adalah sebagai berikut:

a. Probabilitas Kejadian Kejang Berdasarkan Jenis Kelamin

Nilai probabilitas bersyarat pada variabel kejadian kejang dengan variabel jenis kelamin disajikan pada Tabel 4.10.

Tabel 4.10 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Jenis Kelamin

Jenis Kelamin	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
Laki-laki	0,5652	0,5556
Perempuan	0,4348	0,4444

Pada Tabel 4.10 diketahui bahwa probabilitas penderita epilepsi yang berjenis kelamin laki-laki dan tidak mengalami keja-

dian kejang dalam waktu satu bulan terakhir adalah sebesar 0,5652 dan probabilitas penderita epilepsi laki-laki yang mengalami kejang dengan jenis kelamin laki-laki adalah sebesar 0,5556. Sedangkan untuk probabilitas penderita epilepsi perempuan yang tidak mengalami kejadian kejang dan mengalami kejadian kejang dalam waktu satu bulan terakhir adalah sebesar 0,4348 dan 0,4444. Sehingga dapat dikatakan bahwa penderita epilepsi yang tidak mengalami kejadian kejang dengan jenis kelamin laki-laki memiliki nilai probabilitas yang paling besar

b. Probabilitas Kejadian Kejang Berdasarkan Riwayat Demam Kejang

Nilai probabilitas untuk variabel kejadian kejang dengan variabel independen bersifat kategorik yaitu variabel riwayat demam kejang ditampilkan pada Tabel 4.11.

Tabel 4.11 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Riwayat Demam Kejang

Riwayat Demam Kejang	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
Tidak ada	0,8261	0,5556
Ada	0,1739	0,4444

Nilai probabilitas bersyarat untuk penderita epilepsi yang tidak memiliki riwayat demam disertai kejang dan tidak mengalami kejadian kejang adalah sebesar 0,8261, dimana nilai tersebut lebih besar daripada penderita epilepsi yang memiliki riwayat demam disertai kejang dan tidak mengalami kejang dalam waktu satu bulan terakhir yang hanya memiliki probabilitas sebesar 0,1739. Sedangkan untuk penderita epilepsi yang memiliki riwayat disertai kejang yang tidak maupun mengalami kejadian kejang dalam waktu satu bulan terakhir mempunyai nilai probabilitas bersyarat sebesar 0,5556 dan 0,4444.

c. Probabilitas Kejadian Kejang Berdasarkan Konsumsi Obat Antiepilepsi

Variabel independen lainnya yang bersifat kategorik adalah variabel konsumsi obat antiepilepsi, dimana variabel ini memiliki

dua kategori yaitu tidak mengonsumsi obat antiepilepsi dan mengonsumsi obat antiepilepsi. Nilai probabilitas bersyarat dari masing-masing kategori dalam variabel konsumsi obat antiepilepsi dengan variabel kejadian kejang disajikan pada Tabel 4.12.

Tabel 4.12 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Konsumsi Obat Antiepilepsi

Konsumsi Obat Antiepilepsi	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
Tidak	0,2609	0,2222
Ya	0,7391	0,7778

Berdasarkan Tabel 4.12 dapat diketahui bahwa probabilitas penderita epilepsi yang mengonsumsi obat antiepilepsi dan tidak maupun mengalami kejadian kejang dalam waktu satu bulan terakhir memiliki nilai probabilitas yang lebih besar daripada probabilitas penderita epilepsi yang tidak mengonsumsi obat antiepilepsi yaitu sebesar 0,7391 dan 0,7778.

d. Probabilitas Kejadian Kejang Berdasarkan Pemicu Kejang

Probabilitas dari variabel pemicu kejang yang bersyarat variabel kejadian kejang memiliki nilai untuk masing-masing kategori. Hasil menunjukkan nilai probabilitas dari setiap kategori variabel pemicu kejang dengan variabel kejadian kejang. Agar lebih jelas nilai-nilai tersebut disajikan pada Tabel 4.13

Tabel 4.13 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Pemicu Kejang

Pemicu Kejang	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
Tidak memiliki	0,5217	0,1667
Memiliki	0,4783	0,8333

Probabilitas penderita epilepsi yang tidak memiliki faktor pemicu kejang dan tidak mengalami kejadian kejang dalam waktu satu bulan lebih besar dibandingkan dengan yang memiliki faktor pemicu

kejang yaitu sebesar 0,5217. Berbeda halnya untuk probabilitas penderita epilepsi yang tidak memiliki faktor pemicu kejang dan mengalami kejadian kejang dalam waktu satu bulan yang memiliki nilai probabilitas lebih kecil dibandingkan dengan probabilitas penderita epilepsi yang memiliki faktor pemicu kejang yaitu sebesar 0,1667.

e. Probabilitas Kejadian Kejang Berdasarkan Intensitas Terbangun

Selanjutnya variabel independen lain yang dapat dihitung probabilitas bersyaratnya dengan variabel kejadian kejang adalah variabel intensitas terbangun. Nilai probabilitas penderita epilepsi yang memiliki intensitas terbangun saat malam hari kurang dari sama dengan 1 kali dan tidak mengalami kejang adalah sebesar 0,8696, dimana nilai probabilitas tersebut lebih besar daripada penderita epilepsi yang memiliki intensitas terbangun lebih dari satu kali saat malam hari dan tidak mengalami kejadian kejang dalam waktu satu bulan terakhir. Begitu-pula untuk probabilitas penderita epilepsi yang memiliki intensitas terbangun saat malam hari kurang dari atau sama dengan satu kali dan mengalami kejadian kejang yang memiliki nilai probabilitas yang juga lebih besar dibandingkan penderita epilepsi yang memiliki intensitas terbangun lebih dari satu kali saat malam hari dan mengalami kejadian kejang selama satu bulan terakhir. Nilai probabilitas bersyarat untuk setiap kategori variabel intensitas terbangun dan kejadian kejang dapat dilihat pada Tabel 4.14.

Tabel 4.14 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Intensitas Terbangun

Intensitas Terbangun	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
≤ 1 kali	0,8696	0,7778
> 1 kali	0,1304	0,2222

f. Probabilitas Kejadian Kejang Berdasarkan Pemakaian AC

Nilai probabilitas bersyarat variabel pemakaian AC dengan variabel kejadian kejang disajikan pada Tabel 4.15.

Tabel 4.15 Nilai Probabilitas Bersyarat Variabel Kejadian Kejang Berdasarkan Pemakaian AC

Pemakaian AC	Kejadian Kejang	
	Tidak Mengalami Kejang	Mengalami Kejang
Tidak memakai AC	0,6957	0,5556
Memakai AC	0,3043	0,4444

Tabel 4.15 menunjukkan bahwa probabilitas penderita epilepsi yang tidak memakai AC dan tidak mengalami kejadian kejang dalam waktu satu bulan adalah sebesar 0,6957. Nilai probabilitas tersebut lebih besar daripada nilai probabilitas penderita epilepsi yang tidak memakai AC dan mengalami kejang yaitu sebesar 0,5556. Berbeda halnya untuk penderita epilepsi yang memakai AC di rumah dan tidak mengalami kejadian kejang dalam waktu satu bulan memiliki nilai probabilitas bersyarat yang lebih kecil yaitu sebesar 0,3043 daripada penderita epilepsi yang mengalami kejadian kejang yaitu sebesar 0,4444.

Sedangkan untuk variabel yang bersifat kontinu nilai probabilitas bersyaratnya dihitung menggunakan pendekatan distribusi normal. Perhitungan probabilitas bersyarat dengan pendekatan distribusi normal memerlukan nilai rata-rata dan standar deviasi dari masing-masing variabel independen pada setiap kategori kejadian kejang. Tabel 4.16 menyajikan nilai rata-rata dan standar deviasi pada setiap kategori kejadian kejang dari data *training*.

Tabel 4.16 Nilai Rata-Rata dan Standar Deviasi Variabel Usia dan Usia Terdiagnosis Tiap Kategori Kejadian Kejang

Variabel Independen	Tidak Mengalami Kejang		Mengalami Kejang	
	Mean	StDev	Mean	StDev
Usia	25,5217	14,5223	21,5000	12,5147
Usia terdiagnosis	14,4347	13,2075	13,1667	12,7198

Nilai rata-rata dan standar deviasi dari variabel independen pada Tabel 4.16 selanjutnya digunakan untuk menentukan nilai probabilitas bersyarat untuk setiap kategori kejadian kejang menggunakan persamaan (2.18).

Selanjutnya diberikan simulasi perhitungan data *testing* pertama untuk menentukan kategori kejadian kejang. Tabel 4.17 menunjukkan probabilitas bersyarat dari variabel-variabel independen untuk setiap kategori kejadian kejang, dimana nilai probabilitas bersyarat tersebut diperoleh dari perhitungan probabilitas bersyarat pada data *training* yang telah dijelaskan pada Tabel 4.10-Tabel 4.16.

Tabel 4.17 Probabilitas Bersyarat Tiap Kategori pada Data *Testing* Pertama

Variabel	x_{li}	$P(x_{li} \text{Kat.1})$	$P(x_{li} \text{Kat.2})$
Jenis kelamin	1	0,5652	0,5556
Usia	17	0,0231	0,0299
Usia terdiagnosis	9	0,0277	0,0297
Riwayat demam	2	0,1739	0,4444
Konsumsi obat antiepilepsi	2	0,7391	0,7778
Pemicu kejang	2	0,4783	0,8333
Intensitas terbangun	1	0,1304	0,2222
Pemakaian AC	2	0,3043	0,4444

Setelah mendapatkan nilai probabilitas bersyarat untuk masing-masing variabel independen pada setiap kategori kejadian kejang, langkah selanjutnya adalah menghitung probabilitas akhir. Nilai probabilitas akhir diperoleh dari perkalian nilai probabilitas awal pada Tabel 4.9 dengan seluruh nilai probabilitas bersyarat dari variabel independen untuk setiap kategori kejadian kejang. Hasil perkalian yang didapatkan digunakan untuk menentukan kelompok kategori kejadian kejang dari data *testing*. Nilai probabilitas akhir dihitung untuk setiap kategori kejadian kejang.

$$\begin{aligned}
 P(x_1, x_2, \dots, x_8 | y_1) &= P(x_{11} | y_1) \times \dots \times P(x_{18} | y_1) \times P(y_1) \\
 &= 0,5652 \times \dots \times 0,3043 \times 0,5609 \\
 &= 0,0882 \times 10^{-5}
 \end{aligned}$$

$$\begin{aligned}
 P(x_1, x_2, \dots, x_8 | y_2) &= P(x_{11} | y_2) \times \dots \times P(x_{18} | y_2) \times P(y_2) \\
 &= 0,5556 \times \dots \times 0,4444 \times 0,4391 \\
 &= 1,4033 \times 10^{-5}
 \end{aligned}$$

Hasil perhitungan probabilitas akhir tersebut menunjukkan bahwa penderita epilepsi pada observasi pertama di data *testing* yang memiliki ciri-ciri seperti pada Tabel 4.17 memiliki probabilitas tidak mengalami kejadian kejang dalam waktu satu bula terakhir adalah sebesar $0,0882 \times 10^{-5}$, sedangkan probabilitas penderita epilepsi tersebut mengalami kejadian kejang adalah sebesar $1,4033 \times 10^{-5}$. Penderita epilepsi pada observasi pertama di data *testing* akan diklasifikasikan ke dalam kategori yang memiliki nilai probabilitas akhir terbesar, sehingga penderita epilepsi tersebut diklasifikasikan ke dalam kejadian mengalami kejang. Perhitungan tersebut dilakukan untuk semua data *testing*. Hasil klasifikasi dari data *testing* ditunjukkan pada Tabel 4.18.

Tabel 4.18 *Confusion Matrix* Klasifikasi Naïve Bayes

Observasi	Prediksi	
	Tidak mengalami kejang	Mengalami kejang
Tidak mengalami kejang	5	1
Mengalami kejang	2	2

Tabel 4.18 menunjukkan hasil klasifikasi pada data *testing* dengan menggunakan metode naïve bayes. Selanjutnya dilakukan perhitungan untuk mendapatkan nilai ketepatan klasifikasi berdasarkan nilai AUC, sensitifitas, dan spesifisitas seperti berikut:

$$AUC = \frac{1}{2} \left(\frac{5}{5+1} + \frac{2}{2+2} \right) = 0,667$$

$$\text{Sensitivitas} = \frac{5}{5+1} = 0,833$$

$$\text{Spesifisitas} = \frac{2}{2+2} = 0,5$$

Nilai sensitivitas sebesar 83,3% memiliki arti bahwa model klasifikasi yang dihasilkan dari metode naïve bayes mampu mengklasifikasikan penderita epilepsi yang tidak mengalami kejadian kejang ke dalam kategori tidak mengalami kejadian kejang sebesar 83,3%. Sedangkan nilai spesifisitas sebesar 50%, menunjukkan bahwa penderita epilepsi yang mengalami kejang

dalam waktu satu bulan tepat diklasifikasikan ke dalam kategori mengalami kejadian kejang sebesar 50%. Nilai AUC yang dihasilkan dari pengklasifikasian data testing adalah sebesar 66,7%. Nilai tersebut menunjukkan bahwa klasifikasi yang dilakukan termasuk dalam kategori *poor classification*.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berasarkan analisis yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut:

1. Penderita epilepsi dalam Komunitas ODE yang menjadi responden penelitian didominasi oleh lulusan diploma/sarjana, dan jenis kelamin yang paling banyak adalah laki-laki. Dalam waktu satu bulan terakhir ditemukan kejadian tidak mengalami kejang dari 51 penderita epilepsi lebih banyak daripada kejadian mengalami kejang dengan proporsi sebesar 57% dan 43%.
2. Klasifikasi kejadian kejang penderita epilepsi dalam Komunitas ODE dengan regresi logistik biner menghasilkan ketepatan klasifikasi tertinggi pada *fold* ke-5. Diperoleh empat variabel yang berpengaruh signifikan terhadap kejadian kejang pada penderita epilepsi yaitu variabel usia, usia terdiagnosis, riwayat demam kejang, dan faktor pemicu kejang. Seseorang yang terdiagnosis menderita epilepsi pada usia yang semakin tua memiliki peningkatan risiko sebesar 1,124 kali untuk mengalami kejang dalam waktu satu bulan terakhir, penderita epilepsi yang memiliki riwayat demam disertai kejang cenderung untuk mengalami kejang 6,9 kali daripada yang tidak memiliki riwayat demam disertai kejang, penderita epilepsi yang memiliki faktor pemicu kejang juga memiliki kecenderungan yang lebih tinggi untuk mengalami kejang yaitu sebesar 12,5 kali daripada yang tidak memiliki faktor pemicu kejang, dan semakin mudanya usia penderita epilepsi juga akan meningkatkan risiko mengalami kejadian kejang dalam waktu satu bulan terakhir sebesar 1,116 kali. Klasifikasi pada data *testing* berdasarkan model yang terbentuk menghasilkan nilai sensitivitas sebesar 50%, spesifisitas 100%, dan AUC sebesar 75% dimana nilai tersebut masuk dalam kategori *fair classification*.
3. Klasifikasi kejadian kejang dengan metode naïve bayes menghasilkan ketepatan klasifikasi tertinggi pada *fold* ke-5. Pengklasifikasian dari data *testing* menghasilkan nilai AUC sebesar 66,67%, nilai sensitivitas 83,3% dan nilai spesifisitas

50%. Nilai ketepatan klasifikasi yang dihasilkan dari metode naïve bayes termasuk dalam kategori *poor classification* dan hasilnya tidak lebih baik dari metode regresi logistik biner.

4.4 Saran

Berdasarkan penelitian yang telah dilakukan maka saran yang dapat diberikan untuk penelitian selanjutnya yaitu diperlukan penambahan kriteria terhadap sampel yang akan digunakan seperti mempertimbangkan jenis epilepsi yang diderita oleh responden agar tujuan dari penelitian dapat terjawab dengan lebih tepat. Selain itu, perlu dicobakan metode-metode klasifikasi lain yang dapat menghasilkan akurasi yang lebih tinggi. Sedangkan saran yang dapat diberikan kepada penderita epilepsi khususnya yang memiliki faktor pemicu kejang untuk sebisa mungkin menghindari hal-hal yang dapat menimbulkan serangan kejang agar tidak mengalami kejang secara tiba-tiba dikarenakan faktor yang memiliki risiko paling tinggi untuk penderita mengalami kejang adalah faktor pemicu yang dimiliki oleh masing-masing penderita.

DAFTAR PUSTAKA

- Adelow, C., Ahlbom, A., Feychting, M., Johnsson, F., Schwartzbaum, J., dan Tomson, T. (2006). Epilepsy as a risk factor for cancer. *J Neurol Neurosurg Psychiatry*, 77(6), 784-786.
- Agresti, A. (2013). *Categorical Data Analysis*. New Jersey: John Wiley & Sons, Inc.
- Engel, J., dan Pedley, T. (2008). *Epilepsy a Comprehensive Textbook* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.
- Ghozali, I. (2009). *Aplikasi Analisis Multivariate Lanjutan dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.
- Gorunescu, F. (2011). *Data Mining Concepts, Model and Technique*. Jerman: Springer.
- Gujarati, D. (2003). *Ekonometrika Dasar*. Diterjemahkan oleh Sumarno Zein. Jakarta: Erlangga.
- Han, J., Kamber, M., dan Pei, J. (2012). *Data Mining Concepts and Technique Third Edition*. USA: Elsevier Inc.
- Harsono, B. A. (1999). *Buku Ajar Neurologi Klinis*. Yogyakarta: Gadjah Mada University Press.
- Hauser, W. A. (1994). The Prevalence and Incidence of Convulsive Disorders in Children. *Epilepsia*, 35(2), 1-6.
- Hosmer, D., dan Lemeshow, S. (2000). *Applied Logistik Regression* (2nd ed.). New York: John Wiley & Sons, Inc.
- Huang, J., dan Ling, C. X. (2005). Using AUC and Accuracy in Evaluating Learning Algorithms. *Transactions on Knowledge and Data Engineering*, 17(3), 299-310.
- International League Against Epilepsy (ILAE). (1997). The Epidemiology of The Epilepsies: Future Directions. *Epilepsia*, 38, 614-618.
- James, G., Witten, D., Hastie, T., dan Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. New York: Springer Science+Business Media.

- Kleinbaum, D., dan Klein, M. (2002). *Logistic Regression*. New York: Springer Verlag.
- Knudzen, F. U. (2000). Febril Seizure: Treatment and Prognosis. *Epilepsia*, 41(1), 2-9.
- Kristanto, A. (2017). Epilepsi Bangkitan Umum Tonik-Klonik di UGD RSUP Sanglah Denpasar-Bali. *Intisari Sains Medis*, 8(1), 69-73.
- Lynette, G. S., dan Ingrid, E. S. (2007). Febrile Seizures, Clinical Review. *Pubmed*, 334, 307-311.
- Nisa, R. A. (2015). *Hubungan Antara Kualitas Tidur dengan Frekuensi Kejang Pada Pasien Epilepsi di RSUD Dr. Moewardi*. Surakarta: Fakultas Kedokteran Universitas Muhammadiyah Surakarta.
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: ANDI Yogyakarta.
- Raharjo, T. B. (2007). *Faktor-Faktor Resiko Epilepsi pada Anak di Bawah Usia 6 Tahun*. Semarang: Program Pendidikan Dokter Spesialis Ilmu Penyakit Saraf Universitas Diponegoro.
- Rajagukguk, N. (2015). Perbandingan Metode Klasifikasi Regresi Logistik Biner dan Naive Bayes pada Status Pengguna KB di Kota Tegal Tahun 2014. *Jurnal Gaussian*, 4(2), 365-374.
- Rakers, F., Walther, M., Schiffner, R., Rupprecht, S., Rasche, M., dan Kockler, M. (2017). Weather as a risk factor for epileptic seizures: A case-crossover study. *Epilepsia*, 58(7), 1287-1295.
- Riyadi, S., dan Sukarmin. (2009). *Asuhan Keperawatan pada Anak*. Yogyakarta: Graha Ilmu.
- Rubeinstain. (2007). *Lecture Notes Kedokteran Klinis*. Jakarta: Erlangga.
- Salsabila. (2012). *Kualitas Hidup pada Pasien Epilepsi*. Yogyakarta: Fakultas Ilmu Sosial dan Humaniora Universitas Islam Negeri Sunan Kalijaga.
- Samosir, R. O. (2015). Perbandingan Klasifikasi Metode Regresi Logistik Biner dan Radial Basis Function Neural Network pada Berat Bayi Lahir Rendah. *Jurnal Gaussian*, 4(4), 997-1005.

- Sari, R., Triono, A., dan Sutomo, R. (2017). Gangguan Tidur pada Anak dengan Epilepsi dan Faktor yang Mempengaruhi. *Sari Pediatri*, 19(1), 7-13.
- Silbernagl, S., dan Lang, F. (2006). *Teks & Atlas Berwarna Patofisiologi*. Diterjemahkan oleh Iqbal Mochtar Setiawan. Jakarta: EGC.
- Soetomenggolo, T. S., dan Ismael, S. (1999). *Buku Ajar Neurologi Anak*. Jakarta: BP IDAI.
- Sompa. (2016). *Stigma dan Tantangan Epilepsi Menyambut Hari Epilepsi Sedunia*. Makassar: Graha Pena Fajar Online.
- Suwarno, dan Abdillah, A. (2016). Penerapan Algoritma Bayesian Regularization Backpropagation Untuk Memprediksi Penyakit Diabetes. *Jurnal MIPA*, 39(2), 150-158.
- Tjahjadi, P., Dikot, Y., dan Gunawan, D. (2005). Gambaran Umum Mengenai Epilepsi. In P. Riski, *Faktor-Faktor Risiko Terjadinya Epilepsi pada Anak di RSUD Dr. Moewardi*. Surakarta: Fakultas Kedokteran Universitas Sebelas Maret.
- Tripathi, M., Balamurugan, E., Aggarwal, M., Lamba, A., dan Dang, N. (2013). Perceived trigger factors of seizure in persons with epilepsy. *Seizure*, 22, 743-747.
- Tunnajah, S. (2005). *Kebahagiaan pada Orang Dengan Epilepsi (ODE)*. Malang: Fakultas Psikologi Universitas Islam Negeri Maulana Malik Ibrahim
- Vozikis, A., Goulionis, J. E., dan Nikolakis, D. (2012). Risk factors associated with Epilepsy: A case-control study. *Health Science Journal*, 6(3), 509-517.
- Walpole, R. E. (1995). *Pengantar Metode Statistika*. Diterjemahkan oleh Ir. Bambang Sumantri. Edisi Ketiga. Jakarta: Gramedia Pustaka Utama.
- Wu, X., dan Kumar, V. (2009). *The Top Ten Alhoritms in Data Mining*. New York: CRC Press.
- Xhemali, D., Hinde, C. J., dan Stone, R. G. (2009). Naive Bayes vs. Decission Trees vs. Neural Networks in the Classification of Training Web Pages. *International Journal of Computer Science Issues*, 4(1), 16-23.

Yulinda, Mutya. (2015). *Hubungan Antara Riwayat Cedera Kepala Terhadap Kejadian Epilepsi Bangkitan Umum di Poliklinik Saraf Rumah Sakit Umum Daerah Dokter Soedarso Pontianak*. Pontianak: Fakultas Kedokteran Universitas Tanjungpura

LAMPIRAN

Lampiran 1. Kuisisioner Penelitian



KUISIONER PENELITIAN TUGAS AKHIR
Topik : "Faktor Risiko dan Kualitas Hidup
Penderita Epilepsi"
Program Studi Sarjana Departemen Statistika ITS



Identitas Peneliti

Nama : Fraschiska Rizky R. dan Lianna Dwi R.
Departemen : Statistika
Perguruan Tinggi : Institut Teknologi Sepuluh Nopember

Dengan hormat,

Sehubungan dengan penyusunan Tugas Akhir, kami selaku mahasiswa Departemen Statistika ITS Surabaya, memohon kesediaan Bapak/Ibu/Saudara untuk berkenan menjadi responden dalam penelitian kami mengenai "Faktor Risiko dan Kualitas Hidup Penderita Epilepsi". Kesediaan Bapak/Ibu/Saudara untuk mengisi kuisisioner sangat berarti bagi penyusunan penelitian ini. Peneliti akan menjaga kerahasiaan dari jawaban yang Anda berikan. Atas perhatian Bapak/Ibu/Saudara kami ucapkan terimakasih.

PERHATIKAN

Berilah tanda centang (✓) dan isi pertanyaan di bawah ini sesuai keadaan diri Anda.

A. Identitas Responden		
1	Jenis Kelamin	: <input type="checkbox"/> Laki-laki <input type="checkbox"/> Perempuan
2	Usia	:
3	Agama	: <input type="checkbox"/> Islam <input type="checkbox"/> Kristen <input type="checkbox"/> Katolik <input type="checkbox"/> Hindu <input type="checkbox"/> Budha <input type="checkbox"/> Lainnya,
4	Hobi	: <input type="checkbox"/> Membaca <input type="checkbox"/> Menulis <input type="checkbox"/> <i>Handcrafting</i> <input type="checkbox"/> Olahraga <input type="checkbox"/> lainnya
5	Alamat Rumah	:
6	No. HP	:
B. Faktor Sosio-Ekonomi		
1	Status Pernikahan	: <input type="checkbox"/> Belum Menikah <input type="checkbox"/> Menikah <input type="checkbox"/> Pernah Menikah
2	Pendidikan Terakhir	: <input type="checkbox"/> SD <input type="checkbox"/> SMP <input type="checkbox"/> SMA <input type="checkbox"/> Diploma/Sarjana <input type="checkbox"/> Lainnya
3	Pekerjaan	: <input type="checkbox"/> Wiraswasta <input type="checkbox"/> Swasta <input type="checkbox"/> PNS <input type="checkbox"/> Lainnya,
4	Penghasilan Bulanan	: <input type="checkbox"/> < 1 juta <input type="checkbox"/> 1-4 juta <input type="checkbox"/> > 4 juta

C. Faktor Riwayat Kesehatan			
1	Usia saat pertama kali terdiagnosis epilepsi	:	
2	Penyebab penyakit epilepsi	:	
			<input type="checkbox"/> Tidak diketahui <input type="checkbox"/> Diketahui, sebutkan
3	Riwayat keluarga dengan penyakit epilepsi	:	<input type="checkbox"/> Tidak ada <input type="checkbox"/> Ada
4	Riwayat pernatal	:	<input type="checkbox"/> Tidak Ada <input type="checkbox"/> Lahir Prematur <input type="checkbox"/> Berat Badan Lahir Rendah <input type="checkbox"/> Asfiksia <input type="checkbox"/> Lainnya,
5	Riwayat demam disertai kejang	:	<input type="checkbox"/> Pernah <input type="checkbox"/> Tidak pernah
6	Trauma otak/kepala	:	<input type="checkbox"/> Tidak ada <input type="checkbox"/> Ada
7	Frekuensi serangan kejang dalam 2 tahun terakhir	:	<input type="checkbox"/> Tidak pernah <input type="checkbox"/> Jarang (1-2 kali) <input type="checkbox"/> Sering (> 2 kali)
8	Pemicu serangan kejang	:	<input type="checkbox"/> Tidak ada <input type="checkbox"/> Ada, sebutkan: 1. 2. 3.
9	Tindakan preventif saat pasien menduga akan terjadi serangan (berilah ranking dari 1-3)	:	<input type="checkbox"/> Duduk / berbaring <input type="checkbox"/> Berdoa / Mengingat Tuhan <input type="checkbox"/> <i>Self/healing</i> / memotivasi diri untuk mengingat hal-hal positif
D. Faktor Gaya Hidup			
1	Intensitas melakukan ibadah	:	<input type="checkbox"/> Setiap hari <input type="checkbox"/> Setiap minggu <input type="checkbox"/> Setiap bulan <input type="checkbox"/> Lainnya,
2	Konsumsi alkohol	:	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
3	Konsumsi rokok	:	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
4	Konsumsi obat anti epilepsi	:	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya, Jika ya rutin / tidak
5	Rata-rata lama tidur tiap hari	:	<input type="checkbox"/> < 4 jam <input type="checkbox"/> 4-8 jam <input type="checkbox"/> > 8 jam
6	Dapat mengemudi seorang diri	:	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
7	Jika pergi keluar rumah perlu didampingi	:	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
8	Kebiasaan melakukan olah tubuh	:	<input type="checkbox"/> Berjalan <input type="checkbox"/> Berlari <input type="checkbox"/> Yoga <input type="checkbox"/> Lainnya

9	Intensitas melakukan olah raga selama satu bulan	<input type="checkbox"/> Sering (> 4 kali) <input type="checkbox"/> Jarang (1-4 kali) <input type="checkbox"/> Tidak pernah
10	Makanan/ minuman yang dihindari	<input type="checkbox"/> Tidak ada <input type="checkbox"/> Ada, sebutkan 1. 2. 3.
11	Intensitas terbangun saat malam hari	<input type="checkbox"/> Sering (> 3 kali) <input type="checkbox"/> Jarang (1-3 kali) <input type="checkbox"/> Tidak pernah
12	Kondisi tempat tinggal dan kantor	<input type="checkbox"/> Tidak memakai AC <input type="checkbox"/> Memakai AC
E. Stigma Epilepsi		
1	Apakah Anda sudah berdamai dengan penyakit epilepsi?	<input type="checkbox"/> Sudah <input type="checkbox"/> Belum <input type="checkbox"/> Dalam proses
2	Apakah Anda terbuka tentang kondisi kesehatan Anda kepada orang lain?	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
3	Apakah Anda berpikir orang takut pada Anda karena penyakit epilepsi?	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
4	Apakah Anda berpikir epilepsi dapat menurunkan kecerdasan Anda?	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
5	Apakah Anda mengizinkan anak Anda menikah dengan seseorang dengan epilepsi?	<input type="checkbox"/> Tidak <input type="checkbox"/> Ya
F. Kualitas Hidup Pasien Epilepsi		

1. Secara keseluruhan, bagaimana Anda mengukur kualitas hidup Anda?
 (Lingkari salah satu nomor pada skala di bawah ini)



Kualitas hidup
terbaik

Kualitas hidup
terburuk

Pertanyaan nomor 2-13 merupakan pertanyaan apa yang Anda rasakan selama 4 minggu terakhir, pilih salah satu angka antara 1- 6.

	Setiap waktu	Sangat sering	Sering	Beberapa kali	Kadang	Tidak pernah
2. Apakah anda merasa sangat bersemangat ?	1	2	3	4	5	6
3. Apakah Anda pernah menjadi orang yang sangat gugup?	1	2	3	4	5	6
4. Pernahkah Anda	1	2	3	4	5	6

	merasa terisih sehingga tidak ada yang menghibur Anda?						
5.	Apakah Anda pernah merasa tenang dan damai?	1	2	3	4	5	6
6.	Apakah Anda punya banyak energi?	1	2	3	4	5	6
7.	Pernahkan Anda merasa putus asa dan sedih?	1	2	3	4	5	6
8.	Apakah Anda merasa sangat lelah?	1	2	3	4	5	6
9.	Apakah Anda pernah menjadi seseorang yang bahagia?	1	2	3	4	5	6
10.	Apakah Anda merasa lelah?	1	2	3	4	5	6
11.	Pernahkan Anda merasa khawatir akan mengalami kejang lagi?	1	2	3	4	5	6
		Setiap waktu	Sangat sering	Sering	Beberapa kali	Kadang	Tidak pernah
12.	Apakah Anda mengalami kesulitan dalam berpikir dan mengambil keputusan?	1	2	3	4	5	6
13.	Apakah penyakit epilepsi membatasi anda untuk melakukan kegiatan sosial, seperti berkunjung ke rumah saudara atau teman dekat & sejawat?	1	2	3	4	5	6
14.	Bagaimana kualitas hidup Anda selama 4 minggu terakhir (yaitu, bagaimana keadaan Anda)? (<i>lingkari salah satu angka di bawah ini</i>)						
		1	2	3	4	5	
		Sangat baik	Cukup baik	Sedang	Cukup buruk	Sangat buruk	
			Ya, sangat sering	Ya, kadang-kadang	Sangat jarang	Tidak sama sekali	
15.	Selama 1 bulan terakhir, apakah Anda mengalami masalah dengan ingatan, seperti sering lupa?	1	2	3	4		
		Setiap waktu	Sangat sering	Sering	Beberapa kali	Kadang-kadang	Tidak pernah

16.	Kesulitan mengingat hal-hal yang dikatakan orang kepada Anda	1	2	3	4	5	6
17.	Kesulitan konsentrasi saat membaca	1	2	3	4	5	6
18.	Kesulitan konsentrasi saat melakukan kegiatan/ pekerjaan	1	2	3	4	5	6

Pertanyaan selanjutnya berkaitan kegiatan yang dilakukan selama 4 minggu terakhir.

		Sangat sering	Sering	Kadang-kadang	Sangat jarang	Tidak sama sekali
19.	Aktivitas di waktu luang (seperti melakukan hobi atau pergi keluar)	1	2	3	4	5
20.	Mengemudi	1	2	3	4	5

Pertanyaan tentang perasaan saat terjadi kejang

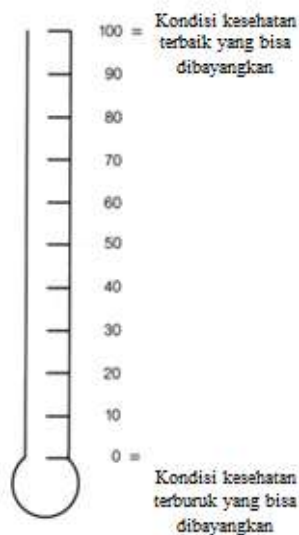
		Sangat takut	Agak takut	Tidak terlalu takut	Tidak takut sama sekali
21.	Seberapa takut anda akan mengalami kejang bulan depan?	1	2	3	4
		Sangat khawatir	Sesekali khawatir	Tidak khawatir sama sekali	
22.	Apakah Anda khawatir akan menyakiti diri sendiri saat kejang	1	2	3	
		Sangat khawatir	Agak khawatir	Tidak terlalu khawatir	Tidak khawatir sama sekali
23.	Seberapa khawatir Anda tentang rasa malu atau masalah sosial lainnya akibat kejang pada bulan berikutnya?	1	2	3	4
		Sangat khawatir	Agak khawatir	Tidak terlalu khawatir	Tidak khawatir sama sekali
24.	Seberapa khawatir Anda tentang efek samping pemakaian obat jangka panjang dalam waktu lama?	1	2	3	4

Untuk masing-masing masalah ini, lingkari satu angka untuk seberapa banyak mereka mengganggu Anda pada skala 1 sampai 5 di mana 1 = sama sekali tidak merepotkan, dan 5 = sangat merepotkan.

		Tidak merepotkan sama sekali	1	2	3	4	Sangat merepotkan
25.	Kejang		1	2	3	4	5

26.	Kesulitan mengingat	1	2	3	4	5
27.	Keterbatasan dalam bekerja	1	2	3	4	5
28.	Keterbatasan dalam bersosial	1	2	3	4	5
29.	Efek fisik dari obat antiepilepsi	1	2	3	4	5
30.	Efek mental dari obat antiepilepsi	1	2	3	4	5

31. Secara keseluruhan, seberapa baik atau buruk menurut Anda tentang kesehatan Anda? Silakan pertimbangkan epilepsi Anda sebagai bagian dari kesehatan Anda ketika Anda menjawab pertanyaan ini.



Lampiran 2. Data Pengujian

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
2	1	27	22	1	1	1	1	1
1	1	23	12	1	2	2	2	1
1	1	40	15	1	1	2	2	1
1	1	39	35	1	2	1	1	1
2	1	29	8	2	2	2	1	1
1	2	29	16	1	2	2	1	1
2	1	26	9	2	1	2	1	1
2	2	20	17	2	2	2	2	1
2	2	38	17	1	2	2	1	2
2	1	27	24	1	2	2	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	2	25	16	1	2	2	1	1
2	2	43	42	1	2	1	1	2
1	2	45	11	2	1	2	1	1
2	1	37	35	1	1	2	1	1
2	2	31	29	2	2	2	1	2

Keterangan:

- y : Kejadian kejang
 x_1 : Jenis kelamin
 x_2 : Usia
 x_3 : Usia terdiagnosis
 x_4 : Riwayat demam disertai kejang
 x_5 : Konsumsi obat antiepilepsi
 x_6 : Pemicu kejang
 x_7 : Intensitas terbangun saat malam hari
 x_8 : Pemakaian AC

Lampiran 3. Syntax k-Fold Cross Validation Regresi Logistik Biner

```

data=read.csv("E:/Data.csv", header=TRUE, sep=";")
data$Kejang=as.factor(data$Kejang)
data$Sex=as.factor(data$Sex)
data$Usia=as.numeric(as.character(data$Usia))
data$Usia.Terdiagnosis=as.numeric(as.character(data$Usia.Terdiagnosis))
data$Demam=as.factor(data$Demam)
data$Konsumsi.Obat=as.factor(data$Konsumsi.Obat)
data$Pemicu=as.factor(data$Pemicu)
data$Terbangun=as.factor(data$Terbangun)
data$AC=as.factor(data$AC)
str(data)
head(data)

library(MXM)
r=5
fold=generatefolds(data$Kejang, nfolds=r, stratified=TRUE, seed = 12345)
akurasi=rep(0,r)
sensi=rep(0,r)
speci=rep(0,r)
auc=rep(0,r)
library(caret)

for(i in 1:r)
{
  train=data[-fold[[i]],]
  test=data[fold[[i]],]
  model=glm(Kejang~., data=train, family=binomial(link="logit"))
  model1=step(model, direction="backward")
  predtest=as.factor(ifelse(predict(model1,test,type="response")>0.5,2,1))
  cm=confusionMatrix(predtest, test$Kejang)
  akurasi[i]=cm$overall['Accuracy']
  sensi[i]=cm$byClass['Sensitivity']
  speci[i]=cm$byClass['Specificity']
  auc[i]=0.5*(sensi[i]+speci[i])
  print(train)
  print(test)
}

```

Lampiran 4. Syntax *k-Fold Cross Validation* Naïve Bayes Classification

```

data=read.csv("e:/data.csv",header=TRUE,sep=";")
data$Kejang=as.factor(data$Kejang)
data$Sex=as.factor(data$Sex)
data$Usia=as.factor(as.character(data$Usia))
data$Usia.Terdiagnosis=as.factor(as.character(data$Usia.Terdiagnosis))
data$Demam=as.factor(data$Demam)
data$Konsumsi.Obat=as.factor(data$Konsumsi.Obat)
data$Pemicu=as.factor(data$Pemicu)
data$Terbangun=as.factor(data$Terbangun)
data$AC=as.factor(data$AC)
str(data)
head(data)

library(MXM)
r=5
fold=generatefolds(data$Kejang, nfolds=r, stratified=T, seed = 12345)
akurasi=rep(0,r)
sensi=rep(0,r)
speci=rep(0,r)
auc=rep(0,r)

library(e1071)
library(Hmisc)
library(naivebayes)
library(caret)
library(rminer)
for(i in 1:r)
{
  train=data[-fold[[i]],]
  test=data[fold[[i]],]
  model=naiveBayes(Kejang~., data=train)
  pretest=predict(model, test)
  cm <- confusionMatrix(pretest, test$Kejang, positive="1",
dnn=c("Prediction","TRUE"))
  akurasi[i] <- cm$overall['Accuracy']
  sensi[i] <- cm$byClass['Sensitivity']
  speci[i] <- cm$byClass['Specificity']
  auc[i] <- 0.5*(sensi[i]+speci[i])
  print(train)
  print(test)
}

```

Lampiran 5. Tabel Kontingensi Antara Variabel Kejadian Kejang dengan Variabel Prediktor yang Bersifat Kategorik

a. Kejadian kejang vs Jenis kelamin

Crosstab

		Kejang		Total
		1,00	2,00	
Sex 1,00	Count	11	10	21
	Expected Count	12,1	8,9	21,0
2,00	Count	12	7	19
	Expected Count	10,9	8,1	19,0
Total	Count	23	17	40
	Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,474 ^a	1	,491		
Continuity Correction ^b	,136	1	,713		
Likelihood Ratio	,476	1	,490		
Fisher's Exact Test				,538	,357
Linear-by-Linear Association	,462	1	,497		
N of Valid Cases	40				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 8,08.

b. Computed only for a 2x2 table

b. Kejadian kejang vs Demam

Crosstab

		Kejang		Total
		1,00	2,00	
Demam 1,00	Count	19	8	27
	Expected Count	15,5	11,5	27,0
2,00	Count	4	9	13
	Expected Count	7,5	5,5	13,0
Total	Count	23	17	40
	Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5,631 ^a	1	,018		
Continuity Correction ^b	4,127	1	,042		
Likelihood Ratio	5,685	1	,017		
Fisher's Exact Test				,038	,021
Linear-by-Linear Association	5,490	1	,019		
N of Valid Cases	40				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,53.

c. Kejadian kejang vs Konsumsi obat

Crosstab

		Kejang		Total	
		1,00	2,00		
Konsumsi_Obat	1,00	Count	5	3	8
		Expected Count	4,6	3,4	8,0
	2,00	Count	18	14	32
		Expected Count	18,4	13,6	32,0
Total		Count	23	17	40
		Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,102 ^a	1	,749		
Continity Correction ^b	,000	1	1,000		
Likelihood Ratio	,103	1	,748		
Fisher's Exact Test				1,000	,537
Linear-by-Linear Association	,100	1	,752		
N of Valid Cases	40				

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 3,40.

d. Kejadian kejang vs Pemicu**Crosstab**

		Kejang		Total
		1,00	2,00	
Pemicu 1,00	Count	8	3	11
	Expected Count	6,3	4,7	11,0
2,00	Count	15	14	29
	Expected Count	16,7	12,3	29,0
Total	Count	23	17	40
	Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,440 ^a	1	,230		
Continuity Correction ^b	,708	1	,400		
Likelihood Ratio	1,489	1	,222		
Fisher's Exact Test				,297	,201
Linear-by-Linear Association	1,404	1	,236		
N of Valid Cases	40				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 4,68.

b. Computed only for a 2x2 table

e. Kejadian kejang vs Terbangun**Crosstab**

		Kejang		Total
		1,00	2,00	
Terbangun 1,00	Count	20	13	33
	Expected Count	19,0	14,0	33,0
2,00	Count	3	4	7
	Expected Count	4,0	3,0	7,0
Total	Count	23	17	40
	Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,744 ^a	1	,388		
Continuity Correction ^b	,195	1	,659		
Likelihood Ratio	,736	1	,391		
Fisher's Exact Test				,432	,326
Linear-by-Linear Association	,726	1	,394		
N of Valid Cases	40				

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,98.

f. Kejadian kejang vs Pemakaian AC

Crosstab

		Kejang		Total
		1,00	2,00	
AC 1,00	Count	17	9	26
	Expected Count	15,0	11,1	26,0
2,00	Count	6	8	14
	Expected Count	8,1	6,0	14,0
Total	Count	23	17	40
	Expected Count	23,0	17,0	40,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,890 ^a	1	,169		
Continity Correction ^b	1,080	1	,299		
Likelihood Ratio	1,885	1	,170		
Fisher's Exact Test				,198	,149
Linear-by-Linear Association	1,843	1	,175		
N of Valid Cases	40				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,95.

Lampiran 6. Output Regresi Logistik Biner Tahap Awal

Logistic Regression Model

lrm(formula = model)

		Model Likelihood Ratio Test		Discrimination Indexes		Rank Discrim. Indexes	
Obs	41	LR chi2	15.23	R2	0.416	C	0.824
1	23	d.f.	8	g	1.985	Dxy	0.647
2	18	Pr(> chi2)	0.0549	gr	7.281	gamma	0.647
max deriv	2e-08			gp	0.330	tau-a	0.327
				Brier	0.166		

	Coef	S.E.	wald z	Pr(> Z)
Intercept	-1.9608	1.4053	-1.40	0.1629
Sex=2	0.5204	0.8627	0.60	0.5463
Usia	-0.1141	0.0586	-1.95	0.0515
Usia.Terdiagnosis	0.1198	0.0655	1.83	0.0674
Demam=2	2.1692	1.1006	1.97	0.0487
Konsumsi.Obat=2	-0.1672	0.9170	-0.18	0.8553
Pemicu=2	2.6702	1.0388	2.57	0.0102
Terbangun=2	0.2507	1.1232	0.22	0.8233
AC=2	0.5741	0.8804	0.65	0.5143

Lampiran 7. Pemilihan Model Terbaik Menggunakan *Backward Elimination*

Start: AIC=59

Kejang ~ Sex + Usia + Usia.Terdiagnosis + Demam + Konsumsi.Obat +
Pemicu + Terbangun + AC

	Df	Deviance	AIC
- Konsumsi.Obat	1	41.034	57.034
- Terbangun	1	41.051	57.051
- Sex	1	41.373	57.373
- AC	1	41.432	57.432
<none>		41.001	59.001
- Usia.Terdiagnosis	1	45.029	61.029
- Usia	1	45.646	61.646
- Demam	1	45.985	61.985
- Pemicu	1	49.521	65.521

Step: AIC=57.03

Kejang ~ Sex + Usia + Usia.Terdiagnosis + Demam + Pemicu + Terbangun + AC

	Df	Deviance	AIC
- Terbangun	1	41.090	55.090
- Sex	1	41.419	55.419
- AC	1	41.438	55.438
<none>		41.034	57.034
- Usia.Terdiagnosis	1	45.046	59.046
- Usia	1	45.721	59.721
- Demam	1	45.997	59.997
- Pemicu	1	49.524	63.524

Step: AIC=55.09

Kejang ~ Sex + Usia + Usia.Terdiagnosis + Demam + Pemicu + AC

	Df	Deviance	AIC
- AC	1	41.461	53.461
- Sex	1	41.502	53.502
<none>		41.090	55.090
- Usia.Terdiagnosis	1	45.115	57.115
- Demam	1	45.999	57.999
- Usia	1	46.028	58.028
- Pemicu	1	50.122	62.122

Step: AIC=53.46

Kejang ~ Sex + Usia + Usia.Terdiagnosis + Demam + Pemicu

	Df	Deviance	AIC
- Sex	1	42.160	52.160
<none>		41.461	53.461
- Usia.Terdiagnosis	1	46.294	56.294
- Demam	1	46.477	56.477
- Usia	1	47.327	57.327
- Pemicu	1	50.522	60.522

Step: AIC=52.16

Kejang ~ Usia + Usia.Terdiagnosis + Demam + Pemicu

	Df	Deviance	AIC
<none>		42.160	52.160
- Usia.Terdiagnosis	1	46.622	54.622
- Demam	1	46.698	54.698
- Usia	1	47.427	55.427
- Pemicu	1	50.816	58.816

Lampiran 8. Analisis Regresi Logistik Biner pada Model Terbaik

Logistic Regression Model

```
lrm(formula = model)
```

		Model Likelihood Ratio Test		Discrimination Indexes		Rank Discrim. Indexes	
Obs	41	LR chi2	14.07	R2	0.389	C	0.816
1	23	d.f.	4	g	1.790	Dxy	0.633
2	18	Pr(> chi2)	0.0071	gr	5.987	gamma	0.633
max deriv	9e-10			gp	0.318	tau-a	0.320
				Brier	0.174		

	Coef	S.E.	wald Z	Pr(> Z)
Intercept	-1.5181	0.9986	-1.52	0.1285
Usia	-0.1097	0.0536	-2.05	0.0408
Usia.Terdiagnosis	0.1171	0.0614	1.91	0.0562
Demam=2	1.9297	0.9889	1.95	0.0510
Pemicu=2	2.5276	0.9649	2.62	0.0088

Hosmer and Lemeshow goodness of fit (GOF) test

```
data: model$y, fitted(model)
```

```
X-squared = 5.7722, df = 8, p-value = 0.6727
```

(Intercept)	Usia	Usia.Terdiagnosis	Demam2	Pemicu2
0.2191309	0.8961120	1.1242808	6.8876073	12.5236815

Lampiran 9. Confusion Matrix Data Testing untuk Regresi Logistik Biner

Confusion Matrix and Statistics

Reference
Prediction 1 2
1 3 0
2 3 4

Accuracy : 0.7
95% CI : (0.3475, 0.9333)
No Information Rate : 0.6
P-Value [Acc > NIR] : 0.3823

Kappa : 0.4444
McNemar's Test P-Value : 0.2482

Sensitivity : 0.5000
Specificity : 1.0000
Pos Pred Value : 1.0000
Neg Pred Value : 0.5714
Prevalence : 0.6000
Detection Rate : 0.3000
Detection Prevalence : 0.3000
Balanced Accuracy : 0.7500

'Positive' Class : 1

Lampiran 10. Output Naïve Bayes dari Data *Training*

```

Naive Bayes Classifier for Discrete Predictors

Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:
Y
      1      2
0.5609756 0.4390244

Conditional probabilities:
Sex
Y      1      2
1 0.5652174 0.4347826
2 0.5555556 0.4444444

Demam
Y      1      2
1 0.8260870 0.1739130
2 0.5555556 0.4444444

Konsumsi.Obat
Y      1      2
1 0.2608696 0.7391304
2 0.2222222 0.7777778

Pemicu
Y      1      2
1 0.5217391 0.4782609
2 0.1666667 0.8333333

Terbangun
Y      1      2
1 0.8695652 0.1304348
2 0.7777778 0.2222222

AC
Y      1      2
1 0.6956522 0.3043478
2 0.5555556 0.4444444

```

Lampiran 11. Confusion Matrix Data Testing untuk Naïve Bayes

Confusion Matrix and Statistics

```
      TRUE
Prediction 1 2
          1 5 2
          2 1 2
```

```
      Accuracy : 0.7
      95% CI   : (0.3475, 0.9333)
      No Information Rate : 0.6
      P-Value [Acc > NIR] : 0.3823

      Kappa : 0.3478
      McNemar's Test P-Value : 1.0000

      Sensitivity : 0.8333
      Specificity : 0.5000
      Pos Pred Value : 0.7143
      Neg Pred Value : 0.6667
      Prevalence : 0.6000
      Detection Rate : 0.5000
      Detection Prevalence : 0.7000
      Balanced Accuracy : 0.6667

      'Positive' Class : 1
```

Lampiran 12. Surat Pernyataan Keaslian Data**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Lianna Dwi Rahmawati

NRP : 0621154000055

menyatakan bahwa data yang digunakan dalam Tugas Akhir-Thesis ini benar-benar merupakan hasil survei pada anggota Komunitas ODE dengan data sebagaimana terlampir.

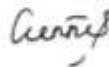
Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data, maka saya siap menerima sanksi sesuai aturan yang berlaku.

Menyetujui,
Ketua Komunitas ODE



Delyma

Surabaya, 14 Juni 2019



Lianna Dwi Rahmawati
NRP. 0621154000055

Mengetahui,
Dosen Pembimbing



Dr. Santi Wulan Purnami
NIP. 19720923 199803 2 001

*(coret yang tidak perlu)

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis memiliki nama lengkap Lianna Dwi Rahmawati atau yang biasa disapa Lianna merupakan anak kedua dari dua bersaudara yang lahir di Madiun, 27 Maret 1997. Putri dari pasangan Bapak Luhur Budi Santoso dan Ibu Tri Rahajoeningtyas ini berdomisili di Madiun dan telah menempuh pendidikan formal di SDN Oro-Oro Ombo Madiun (2003-2009), SMPN 5 Madiun (2009-2012), dan SMAN 2 Madiun (2012-2015). Penulis memilih untuk melanjutkan studi guna menempuh gelar sarjana di Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Pada tahun 2015, penulis dinyatakan lolos SNMPTN sebagai mahasiswa Departemen Statistika ITS. Semasa kuliah penulis aktif di organisasi kemahasiswaan tingkat jurusan yakni Divisi *Statistics Computer Course* yang merupakan Badan Semi Otonom dari Himpunan Mahasiswa Statistika ITS (HIMASTA-ITS), pada periode 2016/2017 sebagai *staff* dalam kemanajeran *Public Relation* dan pada periode 2017/2018 sebagai Sekretaris Divisi. Selain itu penulis turut berpartisipasi dalam kepanitiaan seperti Pekan Raya Statistika, *Indonesia Statistics Conference and Olympiade*, dan ITS OPEN. Selain aktif dalam organisasi dan kepanitiaan, penulis juga mengikuti beberapa perlombaan di bidang statistika dan karya tulis ilmiah. Capaian prestasi penulis antara lain PKM Penelitian Eksakta terdananai pada tahun 2019, finalis *Indonesia Research Competition* pada tahun 2018, finalis *Engineering and Science Festival* pada tahun 2018, dan semifinalis *National Statistics Challenge* pada tahun 2017. Pada Bulan Juli-Agustus 2018 penulis berkesempatan untuk melakukan kerja praktik di Balai Penelitian Tanaman Jeruk dan Buah Subtropika, Batu. Segala kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini dapat dikirimkan melalui *e-mail* liannadwi7@gmail.com