



TUGAS AKHIR - KS184822

**ANALISIS PEMILIHAN KEPUTUSAN PENYESUAIAN
KONSUMSI MAKANAN RUMAH TANGGA SEBAGAI
STRATEGI COPING TERHADAP KEKERINGAN DI
NUSA TENGGARA TIMUR MENGGUNAKAN
REGRESI PROBIT DAN *RANDOM FOREST***

**PRATIWI PENTA ATRIVI
NRP 062117 4500 0009**

**Dosen Pembimbing
Dr. rer. pol. Heri Kuswanto, S.Si., M.Si**

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



TUGAS AKHIR - KS184822

**ANALISIS PEMILIHAN KEPUTUSAN PENYESUAIAN
KONSUMSI MAKANAN RUMAH TANGGA SEBAGAI
STRATEGI COPING TERHADAP RISIKO
KEKERINGAN DI NUSA TENGGARA TIMUR
MENGUNAKAN REGRESI PROBIT DAN *RANDOM
FOREST***

**PRATIWI PENTA ATRIVI
NRP 062117 4500 0009**

**Dosen Pembimbing
Dr. rer. pol. Heri Kuswanto, S.Si., M.Si**

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



FINAL PROJECT - KS184822

**ANALYSIS OF DECISION SELECTION IN
ADJUSTMENT HOUSEHOLD FOOD
CONSUMPTION AS A COPING STRATEGY ON
DROUGHT RISK IN EAST NUSA TENGGARA
USING PROBIT REGRESSION AND RANDOM
FOREST**

**PRATIWI PENTA ATRIVI
NRP 062117 4500 0009**

**Supervisor
Dr. rer. pol. Heri Kuswanto, S.Si., M.Si**

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

LEMBAR PENGESAHAN

ANALISIS PEMILIHAN KEPUTUSAN PENYESUAIAN KONSUMSI MAKANAN RUMAH TANGGA SEBAGAI STRATEGI *COPING* TERHADAP RISIKO KEKERNGAN DI NUSA TENGGARA TIMUR MENGGUNAKAN REGRESI PROBIT DAN *RANDOM FOREST*

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 :

Pratiwi Penta Atrivi
NRP. 062117 4500 0009

Disetujui oleh Pembimbing:

Dr. rer. pol. Heri Kuswanto, S.Si., M.Si
NIP. 19820326 000312 1 004



SURABAYA, JULI 2019

Halaman ini sengaja dikosongkan

ANALISIS PEMILIHAN KEPUTUSAN PENYESUAIAN KONSUMSI MAKANAN RUMAH TANGGA SEBAGAI STRATEGI *COPING* TERHADAP RISIKO KEKERINGAN DI NUSA TENGGARA TIMUR MENGGUNAKAN REGRESI PROBIT DAN *RANDOM FOREST*

Nama : Pratiwi Penta Atrivi
NRP : 062117 4500 0009
Departemen : Statistika-FMKSD-ITS
Pembimbing : Dr. rer. pol. Heri Kuswanto, S.Si., M.Si

Abstrak

Kekeringan adalah salah satu bencana di Indonesia yang berisiko menimbulkan kerugian secara materil maupun non-materil. Kekeringan di Indonesia memiliki risiko tinggi terhadap sektor pertanian yang kemungkinan besar akan berimbas dengan pemenuhan kebutuhan masyarakat (makanan) apabila produktivitas lahan pertanian berkurang, hingga menyebabkan kondisi terburuk yaitu krisis pangan. Salah satu daerah yang memiliki jumlah kejadian kekeringan terbanyak adalah di Nusa Tenggara Timur (NTT) yaitu mencapai 22 kejadian dari total 129 kejadian yang ada di Indonesia pada tahun 2018. Data yang digunakan pada penelitian ini adalah data sekunder hasil survei rumah tangga kasus kekeringan tahun 2018 di NTT. Hasil survei rumah tangga di NTT, diketahui bahwa ternyata lebih dari 60% rumah tangga memilih untuk melakukan penyesuaian konsumsi makanan (dengan mengurangi kualitas dan proporsi makan per hari) sebagai strategi coping terhadap risiko kekeringan. Regresi probit biner menunjukkan rumah tangga dalam memilih keputusan untuk melakukan penyesuaian konsumsi makanan sebagai strategi coping dipengaruhi oleh kepemilikan lahan pribadi, pendapatan rumah tangga, sumber mata pencaharian utama bagi keluarga dan kepemilikan akses tanah dengan ketepatan klasifikasi sebesar 66,3%. Pada analisis random forest menunjukkan ketepatan klasifikasi sebesar 76%, dengan variabel yang memiliki peranan paling penting adalah kepemilikan lahan pribadi.

Kata Kunci: Kekeringan, Random Forest, Regresi Probit, Strategi Coping

Halaman ini sengaja dikosongkan

ANALYSIS OF DECISION SELECTION IN ADJUSTMENT HOUSEHOLD FOOD CONSUMPTION AS A COPING STRATEGY ON DROUGHT RISK IN EAST NUSA TENGGARA USING PROBIT REGRESSION AND RANDOM FOREST

Name : Pratiwi Penta Atrivi
Student Number : 062117 4500 0009
Department : Statistics
Supervisor : Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.

Abstract

Drought is one of the disasters in Indonesia that caused material and non-material losses. Drought in Indonesia has a high risk in the agricultural sector which is likely to have an impact on meeting the needs of the community food if the productivity of agricultural land is reduced, leading to food insecurity. One of the several drought areas is East Nusa Tenggara (NTT), which experienced 22 drought events out of 29 events in Indonesia in 2018. The data used in this study is secondary data from a drought adjustment survey in NTT in 2018. The results showed that more than 60% of households choose to adjust food consumption (by reducing the quality and proportion of food per day) as a coping strategy for the risk of drought. Binary probit regression shows that the households a decision to adjust food consumption as coping strategies are influenced by private land ownership, household income, the main source of livelihood for the family and ownership of land access with classification accuracy of 66.3%. With random forest analysis, the classification accuracy was 76%, with the most important role being private land ownership.

Keywords: *Coping Strategy, Drought, Probit Regression, Random Forest*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah memberikan rahmat, taufiq, dan hidayah-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**Analisis Pemilihan Keputusan Penyesuaian Konsumsi Makanan Rumah Tangga Sebagai Strategi *Coping* Terhadap Risiko Kekeringan di NTT Menggunakan Regresi Probit dan *Random Forest***”. Penyusunan Tugas Akhir ini dapat terselesaikan dengan baik dan lancar karena tidak lepas dari dukungan berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Orang tua dan keluarga tercinta yang selalu memberikan motivasi, dukungan dan doa dalam proses Tugas Akhir ini.
2. Bapak Dr. rer. pol. Heri Kuswanto, S.Si., M.Si. selaku dosen wali dan dosen pembimbing yang telah membimbing, mengarahkan, dan memberikan dukungan dalam menyelesaikan Tugas Akhir ini.
3. Ibu Dr. Vita Ratnasari, S.Si., M.Si dan Dr. R. Mohammad Atok, S.Si., M.Si. selaku dosen penguji yang telah memberikan saran dan kritik yang membangun untuk menyelesaikan Tugas Akhir ini.
4. Bapak Dr. Suhartono selaku Kepala Departemen Statistika ITS yang telah memberikan fasilitas dan sarana dalam penyusunan Tugas Akhir ini.
5. Ibu Santi Wulan Purnami selaku Kepala Program Studi S1 Departemen Statistika ITS yang telah membimbing dan memotivasi penulis selama menjadi mahasiswa.
6. Seluruh dosen Departemen Statistika ITS yang telah memberikan ilmu selama penulis menempuh pendidikan, beserta seluruh karyawan Departemen Statistika ITS yang telah membantu kelancaran dan kemudahan dalam pelaksanaan kegiatan perkuliahan.

7. Teman-teman yang selalu memberikan motivasi, dukungan dan doa dalam proses Tugas Akhir ini serta semua pihak tidak dapat disebutkan satu persatu 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, Juni 2019

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif.....	7
2.2 <i>Missing Value</i>	7
2.3 Tabel Kontingensi (<i>Cross Tabulation</i>).....	8
2.4 Multikolinieritas.....	8
2.5 Regresi Probit Biner.....	9
2.6 <i>Random Forest</i>	16
2.7 Ketepatan Klasifikasi.....	19
2.8 Kekeringan.....	20
2.9 Strategi <i>Coping</i>	21
BAB III METODOLOGI PENELITIAN	23
3.1 Sumber Data.....	23
3.2 Variabel Penelitian.....	23
3.3 Struktur Data.....	24
3.4 Langkah Analisis.....	25
3.5 Diagram Alir.....	27

BAB IV ANALISIS DAN PEMBAHASAN	29
4.1 <i>Missing Value</i>	29
4.2 Statistika Deskriptif	30
4.3 Multikolinieritas.....	37
4.4 Regresi Probit	39
4.4.1 Analisis Model Awal	39
4.4.2 Analisis Model Akhir.....	40
4.4.3 Pengujian Kesesuaian Model	41
4.4.4 Model Probit	42
4.4.5 Ketepatan Klasifikasi	47
4.5 <i>Random Forest</i>	48
4.5.1 Penentuan <i>Mtry</i> dan <i>Ntree</i> Optimum	48
4.5.2 Pembentukan Model	49
4.5.3 Perhitungan Akurasi.....	50
4.5.4 <i>Importance Variable</i>	51
4.6 Penerapan Kasus	52
BAB V KESIMPULAN DAN SARAN	55
5.1 Kesimpulan	55
5.2 Saran	55
DAFTAR PUSTAKA	57
LAMPIRAN	61

DAFTAR TABEL

	Halaman
Tabel 2.1	Tabel Kontingensi I×J.....8
Tabel 2.2	Ketepatan Klasifikasi19
Tabel 3.1	Variabel Penelitian.....23
Tabel 3.2	Struktur Data.....24
Tabel 4.1	<i>Missing Value</i>29
Tabel 4.2	Statistika Deskriptif31
Tabel 4.3	Deteksi Multikolinieritas38
Tabel 4.4	Uji Parsial Model Awal39
Tabel 4.5	Uji Parsial Model Akhir.....40
Tabel 4.6	Ketepatan Klasifikasi47
Tabel 4.7	Pengujian <i>Mtry</i> Terbaik48
Tabel 4.8	Pengujian Nilai OOB dengan <i>Ntree</i> Berbeda49
Tabel 4.9	Hasil Prediksi Data Latih (<i>Train</i>)49
Tabel 4.10	Hasil Prediksi Data Uji (<i>Test</i>).....50
Tabel 4.11	<i>Importance Variable</i>51

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir	27
Gambar 4.1 Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan.....	30
Gambar 4.2 Jenis Kelamin Kepala Rumah Tangga Berdasarkan Pemilihan Strategi <i>Coping</i>	32
Gambar 4.3 Kepemilikan Lahan Pribadi Berdasarkan Pemilihan Strategi <i>Coping</i>	32
Gambar 4.4 Manajemen Ternak Berdasarkan Pemilihan Strategi <i>Coping</i>	33
Gambar 4.5 Respon Terhadap Ramalan Cuaca Berdasarkan Pemilihan Strategi <i>Coping</i>	34
Gambar 4.6 Tindakan Penyimpanan Air Berdasarkan Pemilihan Strategi <i>Coping</i>	34
Gambar 4.7 Pendapatan Rumah Tangga Berdasarkan Pemilihan Strategi <i>Coping</i>	35
Gambar 4.8 Sumber Mata Pencarian Utama Bagi Keluarga Berdasarkan Pemilihan Strategi <i>Coping</i>	36
Gambar 4.9 Kepemilikan Akses Tanah Untuk Bercocok Tanam atau Menggembala Ternak Berdasarkan Pemilihan Strategi <i>Coping</i>	37
Gambar 4.10 Sumber Air untuk Tanaman.....	53

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Surat Pernyataan Pengambilan Data	61
Lampiran 2. Data Faktor-Faktor yang Mempengaruhi Pemilihan Keputusan Penyesuaian konsumsi makanan rumah Tangga Sebagai Strategi <i>Coping</i> Terhadap Risiko Kekeringan di NTT	62
Lampiran 3. <i>Missing Value</i>	63
Lampiran 4. Statistika Deskriptif Data Usia Kepala Rumah Tangga (X_2), Total Aset yang Dimiliki (X_3), Proporsi Anggota Keluarga yang Tidak Aktif Dalam Ekonomi (X_5), dan Total Hewan yang Dimiliki Lima Tahun Terakhir (X_6)	64
Lampiran 5. Tabel Frekuensi Pemilihan Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan (Y)	64
Lampiran 6. Tabel Kontingensi Jenis Kelamin Kepala Rumah Tangga Berdasarkan Pemilihan Strategi <i>Coping</i>	64
Lampiran 7. Tabel Kontingensi Kepemilikan Lahan Pribadi Berdasarkan Pemilihan Strategi <i>Coping</i>	65
Lampiran 8. Tabel Kontingensi Manajemen Ternak Berdasarkan Pemilihan Strategi <i>Coping</i>	65
Lampiran 9. Tabel Kontingensi Respon Rumah Tangga Terhadap Ramalan Cuaca Berdasarkan Pemilihan Strategi <i>Coping</i>	65
Lampiran 10. Tabel Kontingensi Tindakan Penyimpanan Air Berdasarkan Pemilihan Strategi <i>Coping</i> Model.....	66
Lampiran 11. Tabel Kontingensi Pendapatan Rumah Tangga Berdasarkan Pemilihan Strategi <i>Coping</i> Klasifikasi.....	66

Lampiran 12. Tabel Kontingensi Sumber Mata Pencaharian Utama Berdasarkan Pemilihan Strategi <i>Coping</i> Parsial	66
Lampiran 13. Tabel Kontingensi Kepemilikan Akses Tanah Untuk Bercocok Tanam atau Menggembala Ternak Berdasarkan Pemilihan Strategi <i>Coping</i>	67
Lampiran 14. Output minitab Metode Regresi Probit Model Awal.....	67
Lampiran 15. Output minitab Metode Regresi Probit Model Akhir	69
Lampiran 16. Probabilitas Pemilihan Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan	72
Lampiran 17. Tabulasi Silang Klasifikasi Antara Observasi dan Hasil Prediksi Model Regresi Probit Biner	73
Lampiran 18. <i>Syntax Random Forest</i>	73
Lampiran 19. <i>Output Random Forest</i>	76

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia adalah negara kepulauan yang secara geografis dilalui oleh garis khatulistiwa serta berada pada pertemuan tiga lempeng tektonik yaitu lempeng Eurasia, lempeng Pasifik, dan lempeng Indo-Eurasia. Posisi Indonesia tersebut menyebabkan adanya variabilitas iklim yang sangat tinggi. Variabilitas iklim tersebut dapat menimbulkan bencana, salah satunya adalah kekeringan yang berisiko menimbulkan kerugian secara materil maupun non-materil. Secara umum, kekeringan adalah periode masa kering yang lebih lama dari kondisi normal dan menyebabkan ketersediaan air yang jauh di bawah kebutuhan air untuk kebutuhan hidup, pertanian, kegiatan ekonomi maupun lingkungan (Nagarajan, 2009).

Kekeringan di Indonesia selain disebabkan oleh menurunnya curah hujan pada periode yang lama, juga disebabkan oleh ketidakaturan suhu permukaan laut yang terjadi seperti fenomena El nino, positif IOD (*Indian Ocean Dipole*), dan siklus monsun. Pola spasial hubungan antara curah hujan dengan ENSO (*El-Nino Southern Oscillation*) dan IOD menunjukkan bahwa kedua fenomena tersebut mempengaruhi fluktuasi hujan selama musim JJA (Juni-Juli-Agustus) dan SON (September-Oktober-November), sedangkan saat musim monsun DJF (Desember-Januari-Februari) dan MAM (Maret-April-Mei) pengaruh kedua fenomena tersebut tidak jelas khususnya di dalam wilayah Indonesia (As-syakur, et al., 2013). Secara temporal terlihat bahwa ada pergerakan dinamis anatara hubungan ENSO dan IOD dengan curah hujan di Indonesia dimana permulaan pengaruh ENSO dan IOD terjadi pada masa JJA di wilayah barat daya Indonesia dan berakhir pada masa DJF di wilayah timur laut Indonesia.

Kekeringan dapat menimbulkan risiko yang sangat luas, kompleks, baik dalam jangka pendek maupun jangka waktu

panjang hingga setelah berakhirnya kekeringan. Hal tersebut disebabkan karena air yang merupakan kebutuhan pokok dan vital untuk seluruh makhluk hidup yang sumberdaya lainnya tidak dapat menggantikannya. Kekeringan di Indonesia memiliki risiko tinggi terhadap sektor pertanian yang kemungkinan besar akan berimbas dengan pemenuhan kebutuhan masyarakat (makanan) apabila produktivitas lahan pertanian berkurang. Salah satu daerah yang memiliki jumlah kejadian kekeringan terbanyak adalah di Nusa Tenggara Timur (NTT) yaitu mencapai 22 kejadian bencana kekeringan dari total 129 kejadian yang ada di Indonesia pada tahun 2018. Terdapat tujuh kabupaten yang mengalami kekeringan ekstrem di tahun 2018 yaitu antara lain Kabupaten Nagekeo, Kabupaten Ende, Lembata, Sumba Timur, Belu, Rote Ndao, dan Kabupaten Kupang. Kekeringan tersebut mengakibatkan sekitar 865.900 korban (jiwa) yang menderita dan mengungsi (BNPB, 2018).

Di tingkat global, kekeringan masih merupakan satu risiko tersembunyi dan di tingkat lokal, dampak-dampak sosial dan ekonomi yang ditimbulkan kekeringan sangat terkonsentrasi di rumah tangga-rumah tangga miskin pedesaan (UNISDR, 2011). Provinsi Nusa Tenggara Timur (NTT) adalah salah satu provinsi dengan tingkat kerentanan terhadap kerawanan pangan tinggi yaitu lebih dari 50% wilayahnya masuk dalam kategori tersebut. Tingkat kecukupan gizi penduduk NTT masih rendah, tahun 2007-2014 rata-rata hanya sebesar 22,94% dan memiliki pola yang terus meningkat selama tiga tahun terakhir (BKP, 2014).

Dampak kekeringan bervariasi sesuai dengan kelas, usia, etnis dan jenis kelamin, karena faktor-faktor tersebut menentukan kerentanan masyarakat. Berbagai upaya dan manajemen risiko kekeringan dilakukan untuk meminimalkan dampak negatif yang akan terjadi selama kekeringan hingga setelah kekeringan terjadi. Suatu bencana tidak hanya bergantung pada ancaman, kerentanan, dan paparan, tetapi juga pada kapasitas menghadapi (*coping capacity*) serta ketahanan dari elemen berisiko. Kapasitas menghadapi bencana kekeringan meliputi strategi-strategi dan

langkah-langkah untuk bertindak secara langsung atas kerusakan selama kejadian tersebut dengan mengurangi dampaknya dengan tindakan yang efisien untuk menghindari efek yang merusak atau merugikan. Hal ini dapat dilakukan masyarakat dengan mengelola sumber daya yang ada dalam situasi sulit (bencana) yang meliputi tindakan prabencana, saat bencana, dan pascabencana. Mengidentifikasi faktor-faktor risiko dan memahami cara masyarakat menangani dan menghadapi bencana (*coping strategies*) serta beradaptasi dengan lingkungan yang berbahaya dianggap sebagai faktor penentu bagi pengurangan risiko dan pengambilan keputusan di tingkat lokal. Strategi *coping* yang dapat dipilih oleh rumah tangga bermacam-macam. Kesesuaian dan efektivitas strategi tergantung pada informasi, pengetahuan dan keterampilan masing-masing petani dan masyarakat (Pandey & Bhandari, 2009). Di tingkat rumah tangga, pengurangan konsumsi makanan dan perubahan konsumsi makanan adalah strategi penting untuk mengatasi sumber daya yang rendah (Mishra, 2007).

Penelitian sebelumnya menyebutkan, model manajemen siklus kekeringan dapat dianggap sebagai strategi pengurangan risiko bencana yang relevan dalam kasus kekeringan di Kenya, dimana salah satu hasil model probit untuk strategi *coping* di Kenya menggunakan strategi *food consumption adjustments* sebagai indikator risiko bencana dapat menunjukkan bahwa konservasi air memiliki efek positif pada pola gizi rumah tangga dan dalam mengurangi risiko kekeringan (Schilderinck, 2009). Penelitian di Meksiko, mengungkapkan bahwa irigasi dapat mengurangi kerentanan jangka pendek terhadap kekeringan, sedangkan strategi perbaikan dalam efisiensi irigasi telah mengurangi kerentanan kekeringan khususnya bagi pertanian di beberapa daerah untuk jangka panjang (Liverman, 1999). Penelitian di Zimbabwe (berfokus di Lembah Zambezi), menunjukkan berbagai strategi yang digunakan oleh petani kecil untuk beradaptasi serta mengatasi dampak kekeringan antara lain meliputi memproduksi tanaman yang tahan kekeringan, varietas

tanaman diversifikasi, pembelian sereal melalui penjualan aset, sedangkan organisasi non pemerintah memberikan bantuan makanan dan mengumpulkan buah-buahan liar (Mavhura, Manatsa, & Mushore, 2015).

Hasil survei rumah tangga di NTT, diketahui bahwa ternyata lebih dari 60% responden masyarakat memilih untuk melakukan penyesuaian konsumsi makanan (mengurangi jumlah konsumsi makanan ataupun beralih jenis makanan yang dikonsumsi) sebagai strategi *coping* mereka terhadap risiko kekeringan. Strategi tersebut dalam jangka pendek memang berpengaruh dalam mengurangi risiko bencana, seperti hasil analisis yang ditunjukkan pada kasus di Kenya, namun untuk jangka panjang akan menyebabkan malnutrisi parah dan meningkatkan risiko kematian bayi (Schilderinck, 2009). Berdasarkan permasalahan tersebut perlu dilakukan penelitian faktor-faktor yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan sebagai strategi *coping* dalam manajemen risiko kekeringan di NTT menggunakan metode regresi model probit dan *random forest*. Metode regresi model probit digunakan untuk mengetahui faktor apa yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan dari model yang terbentuk, sedangkan dengan metode *random forest* akan diketahui pengaruh faktor yang digunakan berdasarkan urutan pohon dari yang paling tinggi pengaruhnya sampai yang paling rendah. Selain itu, saat ini masih jarang penelitian mengenai strategi adaptasi kekeringan menggunakan *machine learning* seperti *random forest*.

Variabel dependen yang digunakan pada penelitian ini adalah keputusan rumah tangga dalam memilih penyesuaian konsumsi makanan yang bernilai 0 jika tidak memilih dan bernilai 1 jika memilih keputusan untuk melakukan penyesuaian konsumsi makanan. Bentuk penyesuaian konsumsi makanan yang dimaksud adalah dengan mengurangi jumlah konsumsi makanan ataupun dengan beralih jenis makanan yang dikonsumsi, seperti contohnya dari mengkonsumsi beras sebagai makanan

pokok kemudian beralih ke umbi-umbian sebagai pengganti makanan pokok. Hasil analisis yang diperoleh dapat membantu pemerintah setempat dalam merumuskan kebijakan di NTT terhadap risiko kekeringan, khususnya apabila banyak rumah tangga yang memilih keputusan penyesuaian konsumsi makanan dalam jangka panjang yang dikhawatirkan dapat meningkatkan gizi buruk dan jumlah kematian.

1.2 Rumusan Masalah

Provinsi Nusa Tenggara Timur (NTT) adalah salah satu provinsi dengan tingkat kerentanan terhadap kerawanan pangan tinggi sebagai dampak kekeringan yaitu lebih dari 50% wilayahnya masuk dalam kategori tersebut. Tingkat kecukupan gizi penduduk NTT masih rendah, tahun 2007-2014 rata-rata hanya sebesar 22,94% dan memiliki pola yang terus meningkat selama tiga tahun terakhir (BKP, 2014). Berbagai macam strategi perlu dilakukan untuk mengurangi dampak negatif terhadap risiko kekeringan. Hasil survei rumah tangga di NTT, diketahui bahwa ternyata lebih dari 60% responden masyarakat memilih untuk melakukan penyesuaian konsumsi makanan sebagai strategi *coping* mereka terhadap risiko kekeringan. Strategi tersebut dalam jangka pendek memang berpengaruh dalam mengurangi risiko bencana, seperti hasil analisis yang ditunjukkan pada kasus di Kenya, namun untuk jangka panjang akan menyebabkan malnutrisi parah dan meningkatkan risiko kematian bayi. Berdasarkan permasalahan tersebut perlu dilakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT menggunakan metode regresi probit dan *random forest*.

1.3 Batasan Masalah

Batasan masalah yang ada dalam penelitian ini adalah penelitian dilakukan dengan menggunakan data survei rumah tangga kasus kekeringan di Nusa Tenggara Timur tahun 2018.

1.4 Tujuan Penelitian

Menjawab permasalahan yang ada maka tujuan dari penelitian ini adalah

1. Mengetahui hasil analisis faktor-faktor yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT menggunakan metode regresi probit.
2. Mengetahui hasil analisis faktor-faktor yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT menggunakan metode *random forest*.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah mendapatkan hasil analisis terbaik mengenai faktor-faktor yang mempengaruhi pemilihan keputusan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT. Bagi pemerintah setempat, hasil analisis yang diperoleh dapat dijadikan sebagai referensi dalam mengambil kebijakan untuk membantu masyarakat NTT agar strategi *coping* yang dilakukan dapat efektif terhadap risiko kekeringan.

BAB II

TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif merupakan metode dalam statistika yang berkaitan dengan pandataan, pengumpulan, penyajian data, dan penyimpulan hasil pengamatan terhadap seluruh kejadian secara kuantitatif yang dapat dideskripsikan dalam bentuk angka maupun visual dengan menggunakan grafik maupun diagram tertentu. Fungsi statistika deskriptif adalah untuk mengetahui karakteristik suatu situasi tertentu misalkan dengan menghitung nilai rata-rata, median, standar deviasi, nilai maksimum, dan nilai minimum pada suatu objek pengamatan. Secara umum statistika deskriptif dapat diartikan sebagai metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Perlu kiranya dimengerti bahwa statistika deskriptif memberikan informasi hanya mengenai data yang ada dan sama sekali tidak menarik inferensia (Walpole, 1995).

2.2 Missing Value

Missing value (data hilang) merupakan salah satu uji data yang akan diproses dengan berbagai metode statistik multivariat. *Missing value* adalah informasi yang tidak tersedia dari sebuah subjek (kasus). Dalam terminology SPSS, data hilang (*missing value*) adalah adanya sel-sel kosong pada satu atau beberapa variabel. *Missing value* (data hilang) terjadi karena informasi untuk sesuatu tentang objek tidak diberikan, sulit dicari atau memang informasi tersebut tidak ada. Metode imputasi yang paling umum dan mudah untuk digunakan adalah mengganti data hilang (*missing value*) dengan nilai rata-rata atau dengan modus tergantung dari jenis datanya. Pada data numerik digunakan cara mengganti data hilang (*missing value*) dengan nilai rata-rata, sedangkan untuk data kategorik maka digunakan nilai modus. Metode ini menghasilkan penduga rata-rata untuk peubah akan sama dengan nilai yang diimputasikan (Longford, 2005).

2.3 Tabel Kontingensi (*Cross Tabulation*)

Tabel kontingensi merupakan tabel yang berisi data jumlah atau frekuensi atau beberapa klasifikasi (kategori). *Cross tabulation* yaitu suatu metode statistik yang menggambarkan dua atau lebih variabel secara simultan dan hasilnya ditampilkan dalam bentuk tabel yang merefleksikan distribusi bersama dua atau lebih variabel dengan jumlah sel yang terbatas (Agresti, 2002). Secara umum jika memiliki dua variabel A dan B, dimana variabel A terdiri atas I sel, yaitu $A_1, A_2, A_3, \dots, A_i, \dots, A_I$ dan variabel B terdiri atas J sel, yaitu $B_1, B_2, B_3, \dots, B_j, \dots, B_J$ maka akan mempunyai tabel dengan baris sebanyak I dan kolom sebanyak J seperti Tabel 2.1.

Tabel 2.1 Tabel Kontingensi I×J

Variabel A	Variabel B				Total
	1	2	...	J	
1	n_{11}	n_{12}	...	n_{1J}	$n_{1.}$
2	n_{21}	n_{22}	...	n_{2J}	$n_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮
I	n_{I1}	n_{I2}	...	n_{IJ}	$n_{I.}$
Total	$n_{.1}$	$n_{.2}$...	$n_{.J}$	$n_{..}$

keterangan :

- n_{ij} : pengamatan pada sel ke- i, j dengan $i = 1, 2, \dots, I$
 dan $j = 1, 2, \dots, J$
- $n_{i.}$: jumlah pengamatan pada sel ke- i dengan $i = 1, 2, \dots, I$
- $n_{.j}$: jumlah pengamatan pada sel-ke- j dengan $j = 1, 2, \dots, J$
- $n_{..}$: jumlah keseluruhan pengamatan pada sel

2.4 Multikolinieritas

Multikolinieritas adalah adanya hubungan linier yang sempurna di antara beberapa atau semua variabel bebas dalam

model regresi (Gujarati, 2003). Mendeteksi multikolinieritas dapat menggunakan VIF (*Varians Inflation Factor*), apabila nilai VIF lebih besar dari 10 maka menunjukkan adanya kasus multikolinieritas antar variabel prediktor. Perhitungan nilai VIF ditunjukkan pada persamaan (2.1) sebagai berikut.

$$VIF_j = \frac{1}{1 - R_j^2} \quad (2.1)$$

$$R_j^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2.2)$$

dimana R_j^2 adalah koefisien determinasi antara variabel X_j yang diregresikan terhadap variabel prediktor lainnya.

2.5 Regresi Probit Biner

Model regresi ini diperkenalkan pertama kali oleh Chaster Bliss di tahun 1935. Analisis regresi probit adalah analisis yang digunakan untuk melihat hubungan variabel respon yang bersifat kategorik (kualitatif) dan variabel-variabel presiktor yang bersifat kualitatif maupun kuantitatif. Model probit menggunakan normal *Cumulatif Distribution Function* (CDF) untuk menjelaskan fungsi persamaanya karena regresi probit menggunakan pendekatan distribusi normal. Pemodelan regresi probit biner diawali dengan memperhatikan model sebagai berikut (Greene, 2008).

$$y^* = \mathbf{x}^T \boldsymbol{\beta} + \varepsilon \quad (2.3)$$

Variabel y^* merupakan variabel dependen *unobserved* atau variabel laten yang diukur dengan variabel independen atau eksplanatori (Gujarati, 2004). Vektor $\mathbf{x} = [1 \ x_1 \ x_2 \ \dots \ x_p]^T$ dan $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_p]^T$ memiliki ukuran $(p+1) \times 1$, dengan p adalah banyaknya variabel prediktor. *Error* ε diasumsikan berdistribusi normal standar. Untuk membentuk model probit

biner, y^* dikategorikan menjadi dua kategori dengan memberikan nilai *threshold* tertentu. Jika $y^* \leq \gamma$, y dikategorikan 0, sedangkan jika $y^* > \gamma$, y dikategorikan 1. Dengan demikian, model probit untuk $y = 0$ adalah

$$P(y = 0 | \mathbf{x}) = P(y^* \leq \gamma) \quad (2.4)$$

$$P(y = 0 | \mathbf{x}) = P(\mathbf{x}^T \boldsymbol{\beta} + \varepsilon \leq \gamma) \quad (2.5)$$

$$P(y = 0 | \mathbf{x}) = P(\varepsilon \leq \gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.6)$$

$$P(y = 0 | \mathbf{x}) = \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.7)$$

dan model probit untuk $y = 1$ adalah

$$P(y = 1 | \mathbf{x}) = P(y^* > \gamma) \quad (2.8)$$

$$P(y = 1 | \mathbf{x}) = 1 - P(y^* \leq \gamma) \quad (2.9)$$

$$P(y = 1 | \mathbf{x}) = 1 - P(\varepsilon \leq \gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.10)$$

$$P(y = 1 | \mathbf{x}) = 1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.11)$$

dengan $\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) = \Phi(\cdot)$ merupakan CDF dari distribusi normal standar, yaitu

$$\Phi(z) = \int_{-\infty}^z \phi(z) dz \quad (2.12)$$

Berikut merupakan fungsi distribusi probabilitas (*probability distribution function*) dari distribusi normal standar.

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), -\infty < z < \infty \quad (2.13)$$

Interpretasi model regresi probit biner tidak berdasarkan nilai koefisien model akan tetapi menggunakan efek marginal. Efek marginal menunjukkan besarnya pengaruh dari setiap variabel prediktor yang signifikan terhadap probabilitas setiap kategori pada variabel respon yang ditunjukkan pada persamaan (2.14) dan (2.15) sebagai berikut.

$$\frac{\partial \hat{P}(y=0|\mathbf{x})}{\partial \mathbf{x}} = -\boldsymbol{\beta} \phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.14)$$

$$\frac{\partial \hat{P}(y=1|\mathbf{x})}{\partial \mathbf{x}} = \boldsymbol{\beta} \phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \quad (2.15)$$

Model regresi probit ini merupakan GLM (*Generalized Linier Model*) dengan komponen random binomial yang termasuk dalam keluarga eksponensial dengan *link function* model yaitu Φ^{-1} , mengubah probabilitas untuk *z-score* dari distribusi normal standar (Agresti, 2002).

2.5.1 Penaksiran Parameter

Salah satu metode penaksiran parameter yang dapat digunakan adalah dengan *Maximum Likelihood Estimation* (MLE). Metode maksimum likelihood merupakan suatu metode yang memaksimalkan fungsi likelihood. Langkah-langkah untuk memperoleh hasil penaksiran parameter model probit dengan menggunakan metode maksimum likelihood adalah sebagai berikut.

1. Mengambil n sampel random yaitu Y_1, Y_2, \dots, Y_n dimana variabel respon memiliki dua kategori atau biner yang berdistribusi Bernoulli $(1, p)$.
2. Mendapatkan koefisien parameter $(\boldsymbol{\beta})$ dengan dengan membentuk fungsi *likelihood* dari n sampel random.

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left\{ \left[\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \right]^{y_i} \left[1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \right]^{1-y_i} \right\} \quad (2.16)$$

3. Melakukan transformasi \ln terhadap persamaan (2.16) agar lebih mudah dalam memaksimalkan $L(\boldsymbol{\beta})$.

$$\begin{aligned} \ln L(\boldsymbol{\beta}) &= \ln \left[\prod_{i=1}^n \left[\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \right]^{y_i} \left[1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \right]^{1-y_i} \right] \quad (2.17) \\ &= \sum_{i=1}^n \left[y_i \ln \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) + (1 - y_i) \ln \{ 1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \} \right] \end{aligned}$$

4. Memaksimumkan fungsi \ln *likelihood* dengan cara melakukan turunan pertama fungsi $\ln L(\boldsymbol{\beta})$ terhadap $\boldsymbol{\beta}$ dan menyamakannya persamaan (2.16) dengan nol.

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^n \left[y_i \frac{\phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})}{\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} + (1 - y_i) \frac{-\phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})}{1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} \right] x_i \quad (2.18) \\ &= \sum_{i=1}^n x_i \phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \left[\frac{y_i}{\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} + \frac{1 - y_i}{1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} \right] \end{aligned}$$

Sehingga

$$\sum_{i=1}^n x_i \phi(\gamma - \mathbf{x}^T \boldsymbol{\beta}) \left[\frac{y_i}{\Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} + \frac{1 - y_i}{1 - \Phi(\gamma - \mathbf{x}^T \boldsymbol{\beta})} \right] = 0 \quad (2.19)$$

Pada persamaan (2.18) menunjukkan hasil penaksiran parameter $\boldsymbol{\beta}$ dengan menggunakan metode MLE, kemudahan didapatkan fungsi implisit sehingga penaksir parameter $\boldsymbol{\beta}$ tidak dapat langsung didapatkan. Oleh karena itu, digunakan proses iterasi dengan metode Newton-Raphson untuk mendapatkan penaksir parameter $\boldsymbol{\beta}$ dari persamaan yang tidak linier (Agresti, 2002). Secara umum, iterasi ke- $(t + 1)$ metode Newton-Raphson ditunjukkan persamaan (2.19) yang digunakan untuk menaksir parameter $\boldsymbol{\beta}$ dengan $t = 0, 1, 2, \dots$

$$\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^{(t)} - \left(\mathbf{H}^{(t)} \right)^{-1} \mathbf{g}^{(t)} \quad (2.20)$$

Vektor $\boldsymbol{\beta}$ dan \mathbf{g} memiliki ukuran $(p+1) \times 1$. Vektor gradien \mathbf{g} ditunjukkan oleh persamaan (2.21).

$$\mathbf{g} = \left(\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_0} \quad \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_1} \quad \dots \quad \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_p} \right) \quad (2.21)$$

Matriks Hessian (\mathbf{H}) merupakan matriks simetris yang berukuran $(p+1) \times (p+1)$ dan memiliki elemen matriks berupa turunan kedua dari $\ln L(\boldsymbol{\beta})$ yang ditunjukkan persamaan (2.22).

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1^2} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_0} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_1} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_p^2} \end{bmatrix} \quad (2.22)$$

Langkah-langkah iterasi Newton-Raphson adalah sebagai berikut.

1. Menentukan nilai awal $\boldsymbol{\beta}^{(0)}$ saat $t = 0$
2. Menghitung $\mathbf{g}^{(0)}$ dan $\mathbf{H}^{(0)}$ yang besarnya tergantung $\boldsymbol{\beta}^{(0)}$
3. Lakukan iterasi seperti yang ditunjukkan oleh persamaan (2.20) mulai dari $t = 0$
4. Proses iterasi Newton-Raphson akan berhenti, jika hasil iterasi sudah konvergen atau $\|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}\| \leq \varepsilon$ dengan ε merupakan bilangan yang sangat kecil. Penaksir parameter $\boldsymbol{\beta}$ yang didapatkan setelah melakukan proses iterasi hingga iterasi yang terakhir (misalnya, iterasi ke- q) adalah $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(q)}$

2.5.2 Pengujian Parameter

Pengujian signifikansi parameter digunakan untuk menentukan apakah variabel-variabel independen yang terdapat dalam model memiliki hubungan dengan variabel dependennya. Pengujian ini dilakukan dalam bentuk-bentuk berikut ini

a. Uji Serentak

Uji serentak dilakukan untuk menguji keberartian koefisien β secara keseluruhan atau simultan dengan menggunakan *Likelihood Ratio Test* (Hosmer & Lemeshow, 2000). Maka uji serentak dapat dilakukan sebagai berikut:

Hipotesis :

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

(secara bersama-sama variabel independen tidak berpengaruh signifikan terhadap model)

H_1 : Minimal ada satu $\beta_i \neq 0$ dengan $i = 1, 2, \dots, p$ (minimal terdapat satu variabel independen berpengaruh signifikan terhadap model)

Statistik Uji Rasio *Likelihood* :

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n} \right)^{n_1} \left(\frac{n_0}{n} \right)^{n_0}}{\prod_{i=1}^n p_i^{y_i} (1-p_i)^{(1-y_i)}} \right] \quad (2.23)$$

keterangan :

$n_1 = \sum_{i=1}^n y_i$ yaitu banyaknya observasi yang dikategorikan sebagai kategori $y=1$

$n_0 = \sum_{i=1}^n (1 - y_i)$ yaitu banyaknya observasi yang dikategorikan sebagai kategori $y=0$

p_i = nilai probabilitas pada observasi ke- i

Daerah penolakan :

Tolak H_0 apabila nilai $G > \chi_{(db, \alpha)}^2$ atau $p\text{-value} < \alpha$

dengan db adalah banyaknya parameter dalam model tanpa β_0 .

b. Uji Parsial

Uji parsial digunakan untuk menguji keberartian koefisien β secara parsial. Berikut adalah uji parsial yang dapat dilakukan :

Hipotesis :

H_0 : $\beta_i = 0$ (variabel prediktor tidak berpengaruh signifikan terhadap variabel respon)

H_1 : $\beta_i \neq 0$ dengan $i = 1, 2, \dots, p$ (variabel prediktor berpengaruh signifikan terhadap variabel respon)

Statistik Uji *Wald* :

$$W = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (2.24)$$

Daerah penolakan :

Tolak H_0 apabila nilai $|W| \geq Z_{\alpha/2}$ atau $p\text{-value} < \alpha$

2.5.3 Pengujian Kesesuaian Model

Pengujian kesesuaian model digunakan untuk mengetahui apakah model tersebut sesuai atau tidak. Salah satu pengujian kesesuaian model dapat dilakukan menggunakan uji *Pearson*. Uji *pearson* membandingkan nilai dari nilai ekspektasi dan observasi dari setiap sel pada table kontingensi, karena pengelompokan dilakukan berdasarkan estimasi probabilitas. Hipotesis yang digunakan adalah sebagai berikut.

H_0 : Model sesuai (tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi dari model)

H_1 : Model tidak sesuai (terdapat perbedaan antara hasil observasi dengan hasil prediksi dari model)

Statistik Uji *Pearson Chi-Square* :

$$\chi^2 = \sum_i^n r_{pi}^2 \quad (2.25)$$

Daerah penolakan :

Tolak H_0 apabila nilai $\chi^2 > \chi_{db,\alpha}^2$ atau $p\text{-value} < \alpha$, dengan db adalah $n - p - 1$.

2.6 *Random Forest*

Random forest (RF) merupakan pengembangan dari metode *Classification and Regression Tree* (CART) dengan menerapkan metode *bootstrap aggregating* (bagging) dan *random feature selection*. RF merupakan metode yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap *node* dilakukan secara acak. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari *root node*, *internal node*, dan *leaf node* untuk mengambil atribut dan data secara acak sesuai ketentuan. *Root node* merupakan simpul yang terletak paling atas, atau biasa disebut sebagai akar dari pohon keputusan. *Internal node* adalah simpul percabangan, dimana *node* ini mempunyai output minimal dua dan hanya ada satu input. Sedangkan *leaf node* atau *terminal node* merupakan simpul terakhir yang hanya memilih satu input atau tidak mempunyai output (Schouten, Frasincar, & Dekker, 2016).

Pohon keputusan dapat dimulai dengan menghitung nilai *entropy* terlebih dahulu dan nilai *information gain*. Nilai *entropy* adalah Nilai *entropy* dapat dihitung seperti pada persamaan (2.26), dan rumus nilai *information gain* seperti pada persamaan (2.27) sebagai berikut (Han, Kamber, & Pei, 2012).

$$Entropy(S) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (2.26)$$

$$Information\ Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S_A|} Entropy(S_v) \quad (2.27)$$

dimana S adalah himpunan kasus, p_i adalah proporsi nilai S terhadap kelas c , $Values(A)$ adalah semua nilai yang mungkin pada himpunan atribut A , S_v adalah subkelas dari S dengan kelas v yang berhubungan dengan atribut A , dan S_A adalah semua nilai yang sesuai dengan atribut A .

Perhitungan entropi digunakan untuk menentukan pembobotan pada setiap segmen atau kelas dalam perhitungan gain. Proses perhitungan gain digunakan untuk menentukan node dari pohon keputusan. Nilai gain terbesar akan menjadi *node* dan

segmen atau kelas-kelasnya akan menjadi akar (*root*) sedangkan yang memiliki nilai entropi 0 akan menjadi daunnya.

Nilai *information gain* tertinggi dari atribut yang ada dapat dijadikan dasar dalam pemilihan atribut sebagai *node*, baik internal node maupun *root node*. Nilai *gain ratio* merupakan hasil perhitungan dari *information gain* yang dibagi dengan *split information*. Nilai *split information* dapat dilihat pada persamaan (2.28) (Barros, de Carvalho, & Freitas, 2015), sedangkan nilai *gain ratio* dapat dilihat pada persamaan (2.29) (Kotsiantis, 2013).

$$\text{Split Information}(S, A) = \sum_{i=1}^c \left(\frac{|S_i|}{|S|} \right) \log_2 \left(\frac{|S_i|}{|S|} \right) \quad (2.28)$$

$$\text{Gain Ratio}(S, A) = \frac{\text{Information Gain}(S, A)}{\text{Split Information}(S, A)} \quad (2.29)$$

dimana *split information* (S, A) adalah nilai estimasi entropi dari variabel input S yang memiliki kelas c dan $|S_i|/|S|$ adalah probabilitas kelas i dalam atribut.

Pada gugus data yang terdiri atas n amatan dan p peubah penjelas, metode *random forest* dapat dilakukan dengan cara sebagai berikut ini (Breiman & Cutler, 2003).

1. Melakukan penarikan contoh acak berukuran n dengan pemulihan pada gugus data *training* (tahapan *bootstrap*).
2. Dengan menggunakan contoh *bootstrap*, pohon dibangun hingga mencapai ukuran maksimum (tanpa pemangkasan). Setiap simpul, pemilihan pemisah dilakukan dengan secara acak memilih m peubah penjelas, dimana $m < p$, yang selanjutnya pemisah terbaik dipilih berdasarkan m peubah penjelas tersebut (tahapan *random feature selection*).
3. Ulangi langkah 1 dan 2 sebanyak k kali, sehingga terbentuk sebuah hutan yang terdiri atas k pohon.

Metode *random forest* harus menentukan m jumlah variabel prediktor yang diambil secara acak dan k pohon yang akan dibentuk agar mendapatkan hasil yang optimal. Menurut Breiman (1996), nilai k yang disarankan untuk digunakan pada metode

bagging juga dicobakan yaitu $k=50$ sudah memberikan hasil yang memuaskan untuk masalah klasifikasi sedangkan menurut Sutton (2005) $k \geq 100$ cenderung menghasilkan tingkat misklasifikasi yang rendah. Ukuran contoh peubah penjelas (m) saat menggunakan metode *random forest* sangat mempengaruhi korelasi dan kekuatan masing-masing pohon.

Menurut Breiman & Cutler (2003), terdapat tiga cara untuk mendapatkan nilai m sebagai jumlah variabel prediktor yang diambil secara acak dengan nilai p adalah banyak variabel independent (bebas) yaitu :

$$m = \frac{1}{2} \lfloor \sqrt{p} \rfloor \quad (2.30)$$

$$m = \lfloor \sqrt{p} \rfloor \quad (2.31)$$

$$m = 2 \times \lfloor \sqrt{p} \rfloor \quad (2.32)$$

dimana p = total variabel prediktor.

Penggunaan m yang tepat akan menghasilkan *random forest* dengan korelasi antar pohon yang cukup kecil, namun kekuatan setiap pohon cukup besar yang dapat ditunjukkan dengan nilai *error* OOB yang diperoleh bernilai kecil. *Error* OOB dihitung dari proporsi misklasifikasi hasil prediksi *random forest* dari seluruh amatan. Terdapat respons suatu amatan diprediksi dengan melakukan penggabungan hasil prediksi k pohon. Pada masalah klasifikasi dilakukan berdasarkan *majority vote* atau kategori atau kelas yang paling sering muncul sebagai hasil prediksi dari k pohon klasifikasi (Breiman, 2001).

2.6.1 Ukuran Tingkat Kepentingan

MDG (*Mean Decrease Gini*) merupakan salah satu ukuran kepentingan (*importance variable*) peubah penjelas yang dihasilkan oleh metode *random forest*. MDG merupakan ukuran untuk melihat kestabilan tiap variabel. Semakin tinggi nilainya maka akan semakin baik (Breiman, 2000). Misalkan terdapat p peubah penjelas dengan $h=1,2,\dots,p$ maka MDG mengukur tingkat

kepentingan peubah penjelas X_h dengan persamaan (2.33) berikut:

$$MDG_h = \frac{1}{k} \sum_t [d(h,t)I(h,t)] \quad (2.33)$$

dimana :

$d(h,t)$: besar penurunan indeks Gini untuk peubah penjelas X_h

$I(h,t)$: memilah simpul t

k : banyaknya pohon dalam random forest (*ukuran random forest*)

2.7 Ketepatan Klasifikasi

Prosedur klasifikasi adalah suatu evaluasi untuk melihat peluang kesalahan klasifikasi yang dilakukan oleh suatu fungsi klasifikasi. Nilai APER (*Apparent Error Rate*) menunjukkan proporsi observasi yang salah diklasifikasikan oleh fungsi klasifikasi (Johnson & Wichern, 2007). Kebalikannya yaitu akurasi merupakan ukuran yang dapat digunakan untuk mengetahui presentase ketepatan klasifikasi yaitu dengan rumus $1 - \text{APER}$. Tabel ketepatan klasifikasi dapat dilihat pada Tabel 2.2 sebagai berikut.

Tabel 2.2 Ketepatan Klasifikasi

Aktual	Prediksi		Total
	$y=0$	$y=1$	
$y=0$	n_{11}	n_{12}	A
$y=1$	n_{21}	n_{22}	B
Total	C	D	E

keterangan :

n_{11} : Jumlah y_i yang masuk ke dalam kategori $y=0$ yang tepat diklasifikasikan ke kelompok $y=0$

n_{12} : Jumlah y_i yang masuk ke dalam kategori $y=0$ yang diklasifikasikan ke kelompok $y=1$

n_{21} : Jumlah y_i yang masuk ke dalam kategori $y=1$ yang diklasifikasikan ke kelompok $y=0$

- n_{22} : Jumlah y_i yang masuk ke dalam kategori $y=1$ yang tepat diklasifikasikan ke kelompok $y=1$
- A : Jumlah keseluruhan y_i yang masuk ke dalam kategori $y=0$
- B : Jumlah keseluruhan y_i yang masuk ke dalam kategori $y=1$
- C : Jumlah keseluruhan y_i yang tepat diklasifikasikan sebagai kelompok $y=0$
- D : Jumlah keseluruhan y_i yang tepat diklasifikasikan sebagai kelompok $y=1$
- E : Jumlah keseluruhan observasi

Sehingga rumus tingkat akurasi dan APER adalah sebagai berikut.

$$Akurasi = \left(\frac{n_{11} + n_{22}}{E} \right) \quad (2.34)$$

$$APER = \left(\frac{n_{12} + n_{21}}{E} \right) = 1 - Akurasi \quad (2.35)$$

2.8 Kekeringan

Secara umum, kekeringan adalah periode masa kering yang lebih lama dari kondisi normal dan menyebabkan ketersediaan air yang jauh di bawah kebutuhan air (Nagarajan, 2009). Terdapat empat kategori kekeringan antara lain adalah sebagai berikut :

1. Kekeringan Meteorologis

Kekeringan meteorologis berasal dari kurangnya curah hujan dan didasarkan pada tingkat kekeringan relatif terhadap tingkat kekeringan normal atau rata-rata dan lamanya periode kering. Perbandingan ini haruslah bersifat khusus untuk daerah tertentu dan bisa diukur pada musim harian dan bulanan, atau jumlah curah hujan skala waktu tahunan. Kekurangan curah hujan sendiri, tidak selalu menciptakan bahaya kekeringan.

2. Kekeringan Hidrologis

Kekeringan hidrologis mencakup berkurangnya sumber-sumber air seperti sungai, air tanah, danau dan tempat-tempat cadangan air.

3. Kekeringan Pertanian

Kekeringan pertanian merupakan dampak dari kekeringan meteorologi dan hidrologi terhadap produksi tanaman pangan dan ternak. Kekeringan ini terjadi ketika kelembapan tanah tidak mencukupi untuk mempertahankan hasil dan pertumbuhan rata-rata tanaman. Kebutuhan air bagi tanaman, bagaimanapun juga, tergantung pada jenis tanaman, tingkat pertumbuhan dan sarana-sarana tanah. Dampak dari kekeringan pertanian sulit untuk bisa diukur karena rumitnya pertumbuhan tanaman dan kemungkinan adanya faktor-faktor lain yang bisa mengurangi hasil seperti hama, alang-alang, tingkat kesuburan tanah yang rendah dan harga hasil tanaman yang rendah.

4. Kekeringan Sosial Ekonomi

Kekeringan sosial ekonomi berhubungan dengan ketersediaan dan permintaan terhadap barang dan jasa dengan jenis kekeringan meteorologis, hidrologis, serta kekeringan pertanian. Konsep dari kekeringan sosial ekonomi adalah mengenai hubungan antara kekeringan dan aktivitas yang dilakukan manusia. Contohnya seperti, pengelolaan lahan yang tidak baik akan semakin memperburuk dampak-dampak dan kerentanan terhadap kekeringan di masa mendatang.

2.9 Strategi Coping

Strategi *coping* merupakan suatu cara atau metode yang dilakukan tiap individu untuk mengatasi dan mengendalikan situasi atau masalah yang dialami dan dipandang sebagai hambatan, tantangan yang bersifat menyakitkan, serta ancaman yang bersifat merugikan (Kertamuda & H., 2009).

Faktor-faktor yang mempengaruhi strategi *coping* meliputi :

1. Kesehatan fisik
2. Keyakinan atau pandangan positif
3. Keterampilan memecahkan masalah

4. Keterampilan sosial
5. Dukungan sosial
6. Materi

Terdapat dua jenis mekanisme *coping* yang dilakukan individu yaitu *coping* yang berpusat pada masalah (*problem focused form of coping mechanism/direct action*) dan *coping* yang berpusat pada emosi (*emotion focused of coping/palliatif form*) (Stuart & Sundeen, 1991).

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan adalah data sekunder yang diperoleh dari hasil survei “*Drought Cycle Management*” Pusat Studi Kebumian, Bencana dan Perubahan Iklim ITS sebanyak 300 rumah tangga tahun 2018 di Nusa Tenggara Timur. Variabel penelitian ditampilkan pada Tabel 3.1. Dari 300 data yang ada terdiri dari 225 data *training* dan 75 data *testing*.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini antara lain sebagai berikut.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Skala
Strategi <i> coping </i> penyesuaian konsumsi makanan (y)	0 = Tidak 1 = Ya	Nominal
Jenis kelamin kepala rumah tangga (x_1)	0 = Laki-laki 1 = Perempuan	Nominal
Usia kepala rumah tangga (x_2)	Usia dalam tahun	Rasio
Total aset yang dimiliki (x_3)	Jumlah semua barang yang dimiliki di rumah (dalam unit)	Rasio
Kepemilikan lahan pribadi (x_4)	0 = Tidak 1 = Ya	Nominal
Proporsi anggota keluarga yang tidak aktif dalam ekonomi (x_5)	Perbandingan antara jumlah anggota keluarga yang tidak aktif dalam ekonomi dengan jumlah keseluruhan anggota keluarga yang ada	Rasio
Total hewan yang dimiliki lima tahun terakhir (x_6)	Jumlah ternak yang dimiliki selama lima tahun terakhir (sapi, domba, unggas, dll)	Rasio

Tabel 3.1 Variabel Penelitian (lanjutan)

Variabel	Keterangan	Skala
Manajemen ternak (x_7)	0 = Tidak 1 = Ya Meliputi kegiatan menjual dan memotong hewan ternak	Nominal
Respon terhadap ramalan cuaca (x_8)	0 = Tidak 1 = Ya	Nominal
Tindakan penyimpanan air (x_9)	0 = Tidak 1 = Ya	Nominal
Pendapatan rumah tangga (x_{10})	0 = Sumber pendapatan tunggal 1 = Berbagai sumber pendapatan	Nominal
Sumber mata pencaharian utama bagi keluarga (x_{11})	0 = Pastoral 1 = Agro-Pastoral 2 = Bisnis skala kecil 3 = Upah kerja	Nominal
Kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12})	0 = Tidak 1 = Ya	Nominal

3.3 Struktur Data

Berdasarkan variabel-variabel penelitian yang digunakan struktur data pada penelitian ini dapat dilihat pada Tabel 3.2 sebagai berikut.

Tabel 3.2 Struktur Data

Rumah Tangga ke-	Variabel Penelitian					
	Y	X_1	X_2	X_3	...	X_{12}
1	y_1	$x_{1,1}$	$x_{2,1}$	$x_{3,1}$...	$x_{9,1}$
2	y_2	$x_{1,2}$	$x_{2,2}$	$x_{3,2}$...	$x_{9,2}$
3	y_3	$x_{1,3}$	$x_{2,3}$	$x_{3,3}$...	$x_{9,3}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	y_n	$x_{1,n}$	$x_{2,n}$	$x_{3,n}$...	$x_{12,n}$

3.4 Langkah Analisis

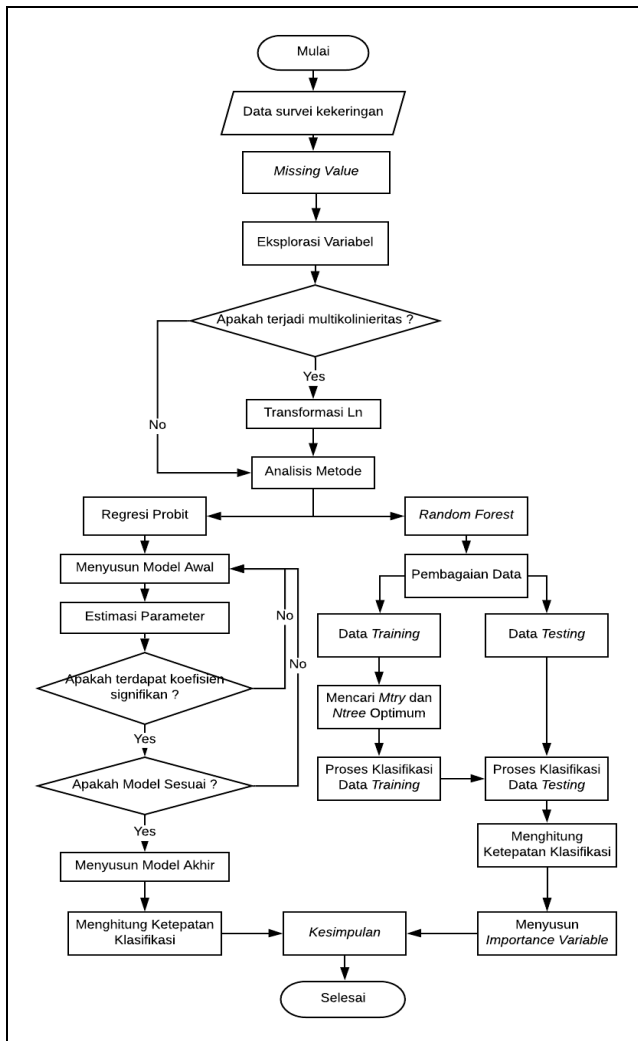
Langkah analisis untuk menggambarkan langkah penelitian menggunakan data hasil survei rumah tangga kasus kekeringan di NTT tahun 2018 ini dapat diuraikan sebagai berikut.

1. Melakukan pengecekan dan penanganan data hilang (*missing value*).
2. Melakukan eksplorasi variabel.
3. Melakukan pengecekan asumsi multikolinieritas, apabila terdapat kasus multikolinieritas maka di atasi dengan transformasi ln.
4. Melakukan analisis pemilihan keputusan penyesuaian konsumsi makanan rumah tangga di NTT menggunakan regresi probit dengan langkah sebagai berikut.
 - a. Pembentukan model awal dari metode regresi probit pada data pemilihan keputusan penyesuaian konsumsi makanan rumah tangga di NTT berdasarkan faktor-faktor yang diduga memiliki pengaruh.
 - b. Melakukan uji rasio *likelihood*, dan uji *wald* pada model awal.
 - c. Pembentukan model akhir dari metode regresi probit pada data pemilihan keputusan penyesuaian konsumsi makanan rumah tangga di NTT berdasarkan faktor-faktor yang signifikan memiliki pengaruh.
 - d. Melakukan uji rasio *likelihood* dan uji *wald* berdasarkan model akhir.
 - e. Melakukan uji kesesuaian model terhadap model regresi probit menggunakan uji *pearson*.
 - f. Menghitung ketepatan klasifikasi model regresi probit biner dengan membandingkan data observasi dengan data hasil prediksi.
5. Melakukan analisis pemilihan keputusan penyesuaian konsumsi makanan rumah tangga di NTT menggunakan *random forest* dengan langkah sebagai berikut.
 - a. Membagi data *training* dan data *testing* dengan presentase 75% dan 25% dari keseluruhan data yang ada.

- b. Melakukan pembentukan *random forest* dengan nilai *mtry* dan *ntree* optimum serta nilai *p* sejumlah dengan variabel prediktor yaitu 12 pada data *training*.
 - c. Melakukan perhitungan akurasi berdasarkan *confusion matrix* hasil proses klasifikasi data latih (*train*) dan data uji (*test*).
 - d. Menyusun *importance variable* berdasarkan data yang memiliki nilai akurasi paling besar.
6. Menarik kesimpulan dan saran.

3.5 Diagram Alir

Langkah analisis penelitian digambarkan dengan menggunakan diagram alir yang ditunjukkan pada Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram Alir

Halaman ini sengaja dikosongkan

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai analisis menggunakan regresi probit dan *random forest*. Namun sebelum melakukan kedua analisis tersebut akan dilakukan pengecekan dan penanganan data hilang (*missing value*), analisis deskriptif untuk mengetahui karakteristik data, dan kemudian deteksi multikolinieritas.

4.1 *Missing Value*

Pengecekan data hilang (*missing value*) digunakan untuk mengetahui apakah pada data pemilihan strategi *coping* penyesuaian konsumsi makanan kasus kekeringan di Nusa Tenggara Timur tahun 2018 terdapat data yang hilang atau tidak. Hasil yang diperoleh dapat dilihat pada Tabel 4.1 berikut.

Tabel 4.1 *Missing Value*

Variabel	<i>Missing Value</i>
Strategi <i>coping</i> penyesuaian konsumsi makanan (y)	0
Jenis kelamin kepala rumah tangga (x_1)	0
Usia kepala rumah tangga (x_2)	2
Total aset yang dimiliki (x_3)	0
Kepemilikan lahan pribadi (x_4)	0
Proporsi anggota keluarga yang tidak aktif dalam ekonomi (x_5)	0
Total hewan yang dimiliki lima tahun terakhir (x_6)	0
Manajemen ternak (x_7)	0
Respon terhadap ramalan cuaca (x_8)	0
Tindakan penyimpanan air (x_9)	1
Pendapatan rumah tangga (x_{10})	0
Sumber mata pencaharian utama bagi keluarga (x_{11})	0
Kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12})	0

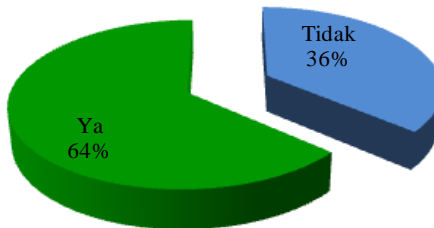
Tabel 4.1 menunjukkan bahwa dari 12 variabel yang digunakan, terdapat dua variabel yang mengandung data hilang

(*missing value*) yaitu variabel usia kepala rumah tangga (x_2) sebanyak 2 data hilang (*missing value*) dan variabel tindakan penyimpanan air (x_9) terdapat 1 data hilang (*missing value*). Penanganan data hilang (*missing value*) ini dapat dengan mengganti data dengan nilai rata-ratanya untuk variabel kontinu dan nilai modus untuk variabel kategorik. Selanjutnya mengganti data variabel usia kepala rumah tangga (x_2) dengan nilai rata-ratanya yaitu 48,748 dan variabel tindakan penyimpanan air (x_9) dengan nilai modus yaitu 1 seperti hasil *output* pada Lampiran 3.

Setelah melakukan penanganan data maka perlu dilakukan pengecekan kembali apakah sudah tidak ada data hilang (*missing value*) dengan melihat hasilnya pada Lampiran 3. Lampiran 3 menunjukkan semua variabel sudah tidak terdapat data hilang (*missing value*).

4.2 Statistika Deskriptif

Setelah data tidak terdapat data hilang (*missing value*) maka dapat dilanjutkan dengan analisis secara deskriptif. Berikut adalah karakteristik data strategi *coping* penyesuaian konsumsi makanan kasus kekeringan di Nusa Tenggara Timur tahun 2018.



Gambar 4.1 Strategi *Coping* Penyesuaian Konsumsi Makanan

Gambar 4.1 menunjukkan dari 300 rumah tangga yang disurvei sebagian besar yaitu 64% (191 rumah tangga) diantaranya memilih strategi *coping* penyesuaian konsumsi makanan dan sisanya sebanyak 36% (109 rumah tangga) tidak memilih strategi *coping* penyesuaian konsumsi makanan sebagai

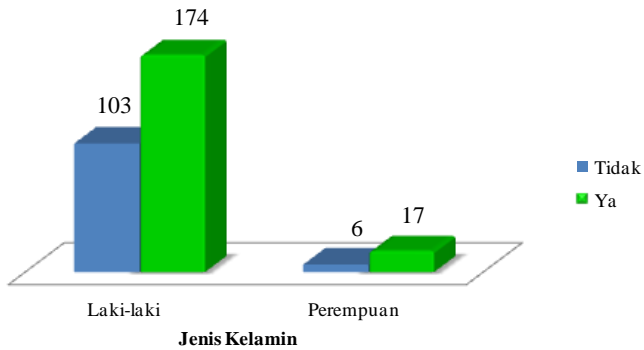
suatu tindakan manajemen resiko kasus kekeringan di NTT pada tahun 2018.

Tabel 4.2 Statistika Deskriptif

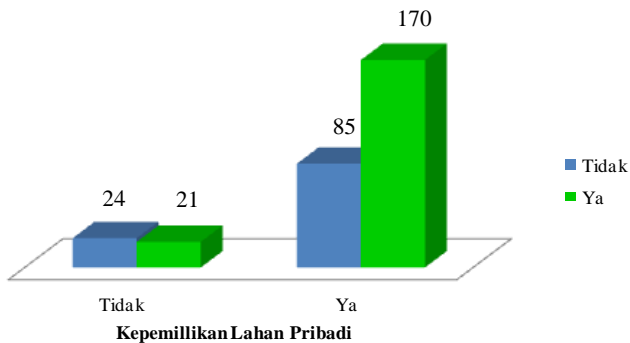
Variabel	Mean	St.Dev	Median	Min	Max
X ₂ = Usia kepala rumah tangga (tahun)	48,748	11,803	48,87	25	88
X ₃ = Total aset yang dimiliki (unit)	3,810	1,508	4	0	10
X ₅ = Proporsi anggota keluarga yang tidak aktif dalam ekonomi	0,398	0,369	0,333	0	3
X ₆ = Total hewan yang dimiliki lima tahun terakhir (ekor)	20,68	26,337	11	0	220

Tabel 4.2 menunjukkan bahwa variabel yang memiliki nilai keragaman paling besar dibanding dengan variabel lainnya adalah variabel total hewan yang dimiliki oleh rumah tangga pada lima tahun terakhir dengan nilai standard deviasi sebesar 26,337. Rumah tangga di Nusa Tenggara Timur tahun 2018 pada lima tahun terakhir memiliki rata-rata sekitar 21 hewan ternak (kerbau, sapi, kambing, domba, unggas dan lain-lain) dengan maksimal hewan ternak yang dimiliki adalah mencapai 220 ekor hewan dan minimum tidak memiliki hewan ternak selama lima tahun terakhir. Selain itu pada Tabel 4.2 menunjukkan 50% rumah tangga memiliki hewan ternak diatas 11 ekor dan 50% lainnya kurang dari 11 ekor hewan ternak ternak (kerbau, sapi, kambing, domba, unggas dan lain-lain) selama lima tahun terakhir.

Dari 300 rumah tangga, mayoritas kepala rumah tangganya adalah laki-laki yaitu sebanyak 277. Dilihat berdasarkan Gambar 4.2 kepala rumah tangga mayoritas kepala keluarga laki-laki di Nusa Tenggara Timur memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan. Hal tersebut dapat dilihat dari total 277 kepala keluarga laki-laki yang ada, 174 diantaranya memilih strategi *coping* tersebut.



Gambar 4.2 Jenis Kelamin Kepala Rumah Tangga Berdasarkan Pemilihan Strategi *Coping*

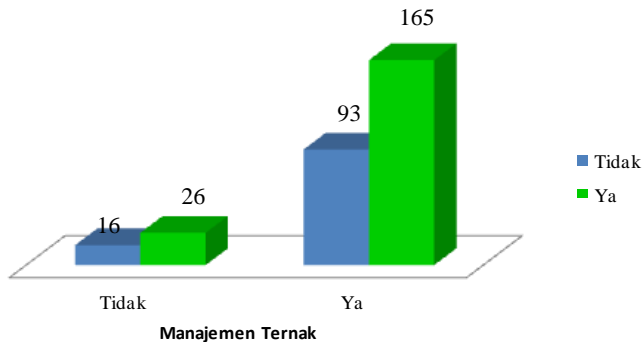


Gambar 4.3 Kepemilikan Lahan Pribadi Berdasarkan Pemilihan Strategi *Coping*

Gambar 4.3 diatas menunjukkan mayoritas rumah tangga di Nusa Tenggara Timur di tahun 2018 sudah memiliki lahan pribadi untuk diolah maupun disewakan lahannya untuk kegiatan ekonomi mereka. Total rumah tangga yang telah memiliki lahan pribadi ada sebanyak 255 dari 300 rumah tangga. Sebanyak 170 rumah tangga diantaranya memilih strategi penyesuaian makanan

sebagai strategi *coping* terhadap risiko kekeringan. Semakin banyak rumah tangga yang memiliki lahan pribadi, maka semakin baik pula tingkat ekonominya karena rumah tangga yang memiliki lahan pribadi tidak perlu membayar uang sewa lahan untuk kegiatan ekonomi.

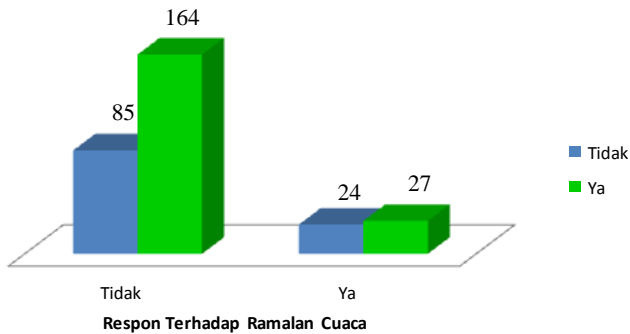
Banyaknya rumah tangga yang melakukan manajemen ternak saat kekeringan dapat dilihat pada Gambar 4.4. Dari total 300 rumah tangga, 258 rumah tangga diantaranya melakukan manajemen ternak sebagai salah satu tindakan dalam menghadapi risiko kekeringan. Gambar 4.4 menunjukkan 165 dari 258 rumah tangga di Nusa Tenggara Timur memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan, sedangkan rumah tangga yang tidak memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan ada sebanyak 93 rumah tangga.



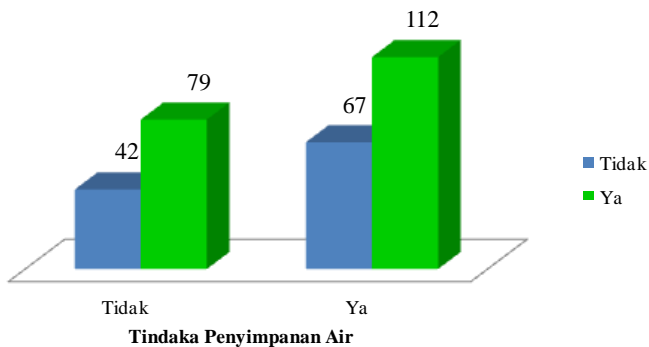
Gambar 4.4 Manajemen Ternak Berdasarkan Pemilihan Strategi *Coping*

Gambar 4.5 dibawah ini menunjukkan bahwa sebagian besar rumah tangga tidak merespon mengenai adanya ramalan cuaca yaitu sebanyak 249 rumah tangga dari 300 rumah tangga. Dari 249 rumah tangga tersebut mayoritas memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan di Nusa Tenggara Timur tahun 2018 yaitu sebanyak

164 rumah tangga dan sisanya 16 rumah tangga tidak memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan. Selain itu, Gambar 4.5 juga menunjukkan hanya sebagian kecil rumah tangga saja yang merespon atau menggunakan ramalan cuaca yaitu ada sebanyak 51 rumah tangga dalam kehidupan sehari-hari mereka.

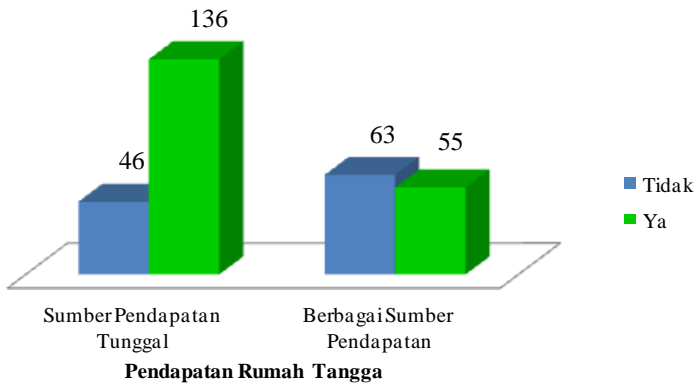


Gambar 4.5 Respon Terhadap Ramalan Cuaca Berdasarkan Pemilihan Strategi *Coping*



Gambar 4.6 Tindakan Penyimpanan Air Berdasarkan Pemilihan Strategi *Coping*

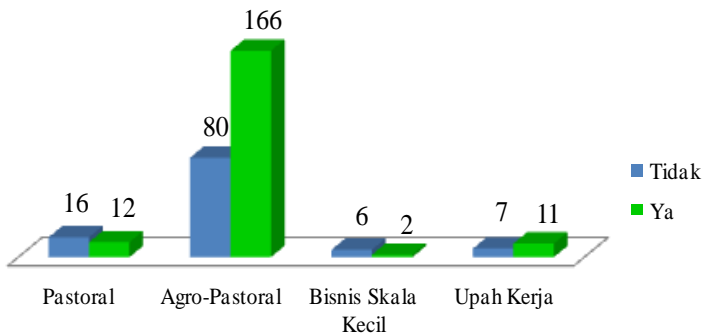
Kasus kekeringan erat kaitannya dengan ketersediaan air, semakin banyak air yang dimiliki rumah tangga untuk kebutuhan sehari-hari maka semakin kecil resiko kekeringan yang akan di terjadi. Salah satu hal yang dapat dilakukan yaitu melakukan suatu penyimpanan air dirumah-rumah. Gambar 4.6 menunjukkan sebagian besar rumah tangga di Nusa Tenggara Timur tahun 2018 telah melakukan tindakan penyimpanan air yaitu terdapat 179 dari total 300 rumah tangga dan sisanya sebanyak 121 rumah tangga tidak melakukan tindakan penyimpanan air dalam menghadapi krisis kekeringan. Jika dilihat dari jumlah rumah tangga yang melakukan penyimpanan air, dari 179 rumah tangga terdapat 112 rumah tangga memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan sedangkan 67 sisanya tidak memilih memilih strategi penyesuaian makanan sebagai strategi *coping*.



Gambar 4.7 Pendapatan Rumah Tangga Berdasarkan Pemilihan Strategi *Coping*

Dari 300 rumah tangga, Gambar 4.7 menunjukkan mayoritas pendapatan rumah tangganya adalah berasal dari sumber pendapatan tunggal yaitu sebanyak 186 rumah tangga. Sedangkan 118 rumah tangga lainnya pendapatannya berasal dari berbagai sumber pendapatan seperti yang terlihat pada Gambar

4.7. Terdapat 136 dari 186 rumah tangga yang masuk dalam kategori sumber pendapatan rumah tangganya adalah tunggal memilih memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan, sedangkan dalam kategori rumah tangga yang memiliki berbagai sumber pendapatan terdapat 63 rumah tangga yang memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan di Nusa Tenggara Timur.



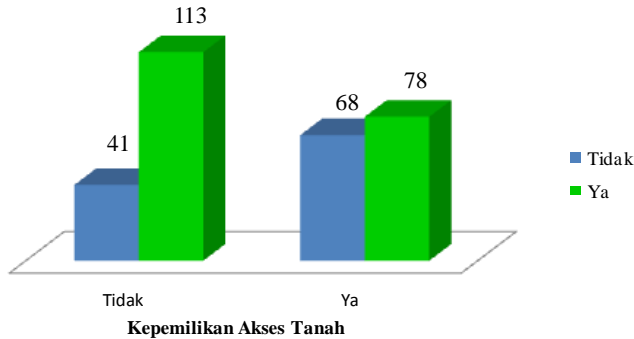
Sumber Mata Pencaharian Utama

Gambar 4.8 Sumber Mata Pencaharian Utama Bagi Keluarga Berdasarkan Pemilihan Strategi *Coping*

Gambar 4.8 menunjukkan jumlah rumah tangga terbanyak dari segi sumber mata pencaharian utama bagi keluarga adalah pada agropastoral dengan 246 rumah tangga yang bermata pencaharian utama dalam bidang tersebut. Sedangkan minoritas rumah tangga mata pencaharian utamanya adalah berasal dari bisnis skala kecil yaitu terdapat 8 rumah tangga.

Selain itu, dapat dilihat pada Gambar 4.8 orang yang bermata pencaharian berasal dari agropastoral 166 rumah tangga memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan. Jumlah tersebut sangat tinggi jika

dibandingkan tiga sumber mata pencaharian utama lainnya di Nusa Tenggara Timur.



Gambar 4.9 Kepemilikan Akses Tanah Untuk Bercocok Tanam atau Menggembala Ternak Berdasarkan Pemilihan Strategi *Coping*

Banyaknya rumah tangga yang telah memiliki akses tanah untuk bercocok tanam atau menggembala ternak di Nusa Tenggara Timur tahun 2018 lebih sedikit dibandingkan dengan yang tidak memiliki akses. Hal tersebut dapat dilihat pada Gambar 4.9 yang menunjukkan masih terdapat 154 rumah tangga yang tidak memiliki akses tanah untuk bercocok tanam maupun menggembala ternak dari total 300 rumah tangga, dan mayoritas diantaranya terdapat 113 rumah tangga memilih strategi penyesuaian makanan sebagai strategi *coping* terhadap risiko kekeringan di Nusa Tenggara Timur tahun 2018.

4.3 Multikolinieritas

Mendeteksi multikolinieritas dapat dilakukan menggunakan VIF (*Varians Inflation Factor*), apabila nilai VIF lebih besar dari 10 maka menunjukkan adanya kasus multikolinieritas antar variabel prediktor seperti persamaan (2.1).

Tabel 4.3 berikut ini adalah nilai VIF untuk setiap variabel prediktor yang diperoleh.

Tabel 4.3 Deteksi Multikolinieritas

Variabel	Keterangan	VIF
Jenis kelamin kepala rumah tangga (x_1)	1=Perempuan	1,04
Usia kepala rumah tangga (x_2)	Usia dalam tahun	1,13
Total aset yang dimiliki (x_3)	Jumlah semua barang yang dimiliki di rumah (dalam unit)	1,27
Kepemilikan lahan pribadi (x_4)	1=Ya	1,13
Proporsi anggota keluarga yang tidak aktif dalam ekonomi (x_5)	Perbandingan antara jumlah anggota keluarga yang tidak aktif dalam ekonomi dengan jumlah keseluruhan anggota keluarga yang ada	1,11
Total hewan yang dimiliki lima tahun terakhir (x_6)	Jumlah ternak yang dimiliki selama lima tahun terakhir (sapi, domba, unggas, dll)	1,27
Manajemen ternak (x_7)	1=Ya	1,16
Respon terhadap ramalan cuaca (x_8)	1=Ya	1,11
Tindakan penyimpanan air (x_9)	1=Ya	1,48
Pendapatan rumah tangga (x_{10})	1=Berbagai Sumber	1,29
Sumber mata pencaharian utama bagi keluarga (x_{11})	1=Agro-Pastoral	2,10
	2=Bisnis Skala Kecil	1,51
	3=Upah Kerja	1,72
Kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12})	1=Ya	1,29

Tabel 4.3 dapat menunjukkan bahwa semua nilai VIF untuk setiap variabel prediktor yang ada pada penelitian ini kurang dari 10, yang artinya pada data ini tidak terjadi kasus multikolinieritas sehingga dapat dilanjutkan pada analisis regresi probit dan *random forest* selanjutnya.

4.4 Regresi Probit

4.4.1 Analisis Model Awal

Pada analisis model ini dilakukan dengan menggunakan semua variabel prediktor yang ada yaitu sebanyak 12 variabel yang diuraikan pada Tabel 3.1 sebelumnya. Hasil didapatkan nilai G sebesar 61,31 dan p -value sebesar 0,000 yang dapat dilihat pada Lampiran 14. Hal ini menunjukkan bahwa dengan menggunakan $\alpha = 0,05$ didapatkan keputusan tolak H_0 , karena nilai G yang didapatkan (61,31) lebih besar dari $\chi^2_{(0.05,14)}$ yaitu 23,684. Sehingga dapat disimpulkan bahwa minimal terdapat satu variabel prediktor yang berpengaruh terhadap model.

Untuk mengetahui variabel mana yang berpengaruh terhadap model, maka perlu dilanjutkan dengan uji parsial dengan hasil sebagai berikut.

Tabel 4.4 Uji Parsial Model Awal

Variabel	Coef	SE Coef	W	P-value	Keputusan
X ₁ (1)	- 0,206	0,338	-0,61	0,542	Gagal Tolak H ₀
X ₂	- 0,0009	0,007	-0,31	0,899	Gagal Tolak H ₀
X ₃	- 0,0843	0,059	-1,42	0,155	Gagal Tolak H ₀
X ₄ (1)	- 0,546	0,228	-2,40	0,016	Tolak H ₀
X ₅	0,253	0,226	1,12	0,262	Gagal Tolak H ₀
X ₆	0,0014	0,0035	0,39	0,693	Gagal Tolak H ₀
X ₇ (1)	0,020	0,245	0,008	0,937	Gagal Tolak H ₀
X ₈ (1)	0,194	0,216	0,90	0,370	Gagal Tolak H ₀
X ₉ (1)	0,538	0,197	2,73	0,006	Tolak H ₀
X ₁₀ (1)	0,682	0,181	3,76	0,000	Tolak H ₀
X ₁₁ (1)	-0,320	0,294	-1,09	0,276	Gagal Tolak H ₀
X ₁₁ (2)	1,012	0,581	1,74	0,081	Gagal Tolak H ₀
X ₁₁ (3)	- 0,324	0,434	-0,75	0,456	Gagal Tolak H ₀
X ₁₂ (1)	0,747	0,181	4,12	0,000	Tolak H ₀

Tabel 4.4 menunjukkan dengan $\alpha = 0,05$ dan $Z_{0,025} = 1,96$ didapatkan keputusan tolak H_0 hanya pada empat variabel dari dua belas variabel prediktor yang ada. Kesimpulan yang diperoleh

yaitu variabel yang berpengaruh terhadap pemilihan keputusan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT antara lain adalah variabel kepemilikan lahan pribadi (x_4), tindakan penyimpanan air (x_9), pendapatan rumah tangga (x_{10}), dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12}).

4.4.2 Analisis Model Akhir

Analisis ini dilakukan dengan hanya menggunakan variabel prediktor yang berpengaruh terhadap pemilihan keputusan penyesuaian konsumsi makanan sebagai strategi *coping* manajemen risiko kekeringan di NTT seperti yang telah diuraikan sebelumnya yaitu kepemilikan lahan pribadi (x_4), tindakan penyimpanan air (x_9), pendapatan rumah tangga (x_{10}), dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12}), dilakukan kembali pengujian signifikansi parameter secara simultan.

Hasil didapatkan nilai G sebesar 48,71 dan p -value sebesar 0,000 yang dapat dilihat pada Lampiran 15. Hal ini menunjukkan bahwa dengan menggunakan $\alpha = 0,05$ didapatkan keputusan tolak H_0 , karena nilai G yang didapatkan (48,715) lebih besar dari $\chi^2_{(0.05,4)}$ yaitu 9,488. Sehingga dapat disimpulkan bahwa minimal terdapat satu variabel prediktor yang berpengaruh terhadap model.

Langkah selanjutnya yaitu melakukan uji parsial untuk mengetahui variabel mana yang signifikan berpengaruh terhadap respon yang dapat dilihat pada hasil sebagai berikut.

Tabel 4.5 Uji Parsial Model Akhir

Variabel	Coef	SE Coef	W	P-value	Keputusan
$X_4(1)$	-0,603	0,213	-2,83	0,005	Tolak H_0
$X_9(1)$	0,512	0,176	2,91	0,004	Tolak H_0
$X_{10}(1)$	0,815	0,164	4,97	0,000	Tolak H_0
$X_{12}(1)$	0,601	0,168	3,59	0,000	Tolak H_0

Tabel 4.5 menunjukkan dengan $\alpha = 0,05$ dan $Z_{0,025} = 1,96$ semua variabel yang ada telah signifikan berpengaruh terhadap

respon karena nilai $|W|$ lebih besar daripada $Z_{0,025} = 1,96$ dan nilai p -value kurang dari $\alpha = 0,05$. Kesimpulan yang diperoleh yaitu variabel yang berpengaruh terhadap pemilihan keputusan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT antara lain adalah variabel kepemilikan lahan pribadi (x_4), tindakan penyimpanan air (x_9), pendapatan rumah tangga (x_{10}), dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12}).

Nilai koefisien yang negatif pada variabel kepemilikan lahan pribadi menunjukkan hubungan yang berbanding terbalik terhadap perubahan strategi *coping* yang dipilih. Rumah tangga yang memiliki lahan pribadi dibanding yang tidak memiliki lahan pribadi cenderung kemungkinan untuk melakukan penyesuaian konsumsi makanan lebih kecil dibanding tidak melakukan penyesuaian konsumsi makanan. Nilai koefisien pendapatan rumah tangga (x_{10}) lebih besar dibanding nilai koefisien variabel tindakan penyimpanan air (x_9) dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12}). Hal ini menunjukkan bahwa pengaruh pendapatan rumah tangga lebih dominan dibandingkan dengan tindakan penyimpanan air dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak terhadap pemilihan keputusan penyimpanan konsumsi makanan sebagai strategi *coping* risiko kekeringan.

4.4.3 Pengujian Kesesuaian Model

Langkah selanjutnya yaitu model perlu dievaluasi apakah model yang terbentuk sudah sesuai atau tidak. Salah satu pengujian untuk mengetahui kesesuaian model adalah dengan menggunakan uji *pearson*. Hasil uji *pearson* yang diperoleh dapat dilihat pada Lampiran 15 yang menunjukkan hasil nilai uji *pearson* sebesar 309,26 dengan p -value sebesar 0,272. Berdasarkan hasil tersebut, maka diperoleh keputusan gagal tolak H_0 karena nilai *pearson* yang diperoleh sebesar 309,26 lebih kecil dari $\chi^2_{(0,05,295)}$ sebesar 336,058. Keputusan yang diperoleh adalah model telah sesuai, antara hasil observasi dengan hasil prediksi.

4.4.4 Model Probit

Model probit terbaik yang diperoleh berdasarkan analisis yang telah dilakukan adalah sebagai berikut:

$$\hat{y} = -0,807 - 0,603x_{4(1)} + 0,512x_{9(1)} + 0,815x_{10(1)} + 0,601x_{12(1)}$$

Setelah mendapatkan model regresi probit terbaik, maka dapat dihitung nilai probabilitas rumah tangga yang tidak memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT dengan $y = 0$ sebagai berikut.

$$\begin{aligned} P(y = 0 | \mathbf{x}) &= \Phi(-0,807 - 0,603x_{4(1)} + 0,512x_{9(1)} + 0,815x_{10(1)} + 0,601x_{12(1)}) \\ &= \Phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= \Phi(0,518) \\ &= 0,698 \end{aligned}$$

Probabilitas rumah tangga di NTT yang memiliki kepemilikan lahan pribadi, melakukan tindakan penyimpanan air, pendapatan rumah tangganya berasal dari berbagai sumber, serta memiliki akses tanah untuk bercocok tanam atau menggembala ternak memiliki probabilitas sebesar 0,698 untuk tidak memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan.

$$\begin{aligned} P(y = 1 | \mathbf{x}) &= 1 - \Phi(-0,807 - 0,603x_{4(1)} + 0,512x_{9(1)} + 0,815x_{10(1)} + 0,601x_{12(1)}) \\ &= 1 - \Phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 1 - \Phi(0,518) \\ &= 1 - 0,698 \\ &= 0,302 \end{aligned}$$

Sedangkan probabilitas rumah tangga di NTT yang memiliki kepemilikan lahan pribadi, melakukan tindakan penyimpanan air, pendapatan rumah tangganya berasal dari berbagai sumber, serta memiliki akses tanah untuk bercocok tanam atau menggembala ternak memiliki probabilitas sebesar 0,302 untuk memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan.

Untuk mengetahui besarnya pengaruh dari setiap variabel prediktor yang signifikan yaitu variabel kepemilikan lahan pribadi (x_4), tindakan penyimpanan air (x_9), pendapatan rumah tangga (x_{10}), dan kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12}) dapat dilihat berdasarkan efek marginal dengan perhitungan sebagai berikut ini.

a. Efek marginal variabel kepemilikan lahan pribadi (X_4)

$$\begin{aligned}\frac{\partial \hat{P}(y=0|\mathbf{x})}{\partial X_4} &= -\beta_4 \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= -\beta_4 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,603 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,603 \phi(0,518) \\ &= 0,603(0,349) \\ &= 0,210\end{aligned}$$

$$\begin{aligned}\frac{\partial \hat{P}(y=1|\mathbf{x})}{\partial X_4} &= \beta_4 \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= \beta_4 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,603 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,603 \phi(0,518) \\ &= -0,603(0,349) \\ &= -0,210\end{aligned}$$

Berdasarkan nilai efek marginal variabel kepemilikan lahan pribadi, menunjukkan bahwa pengaruh kepemilikan lahan pribadi terhadap keputusan rumah tangga di NTT dalam pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan adalah sebesar 0,210. Artinya jika rumah tangga memiliki kepemilikan lahan pribadi maka akan menaikkan kontribusi rumah tangga tersebut masuk dalam kategori yang tidak memilih penyesuaian konsumsi makanan yaitu sebesar 0,210. Sebaliknya, apabila rumah tangga memiliki

kepemilikan lahan pribadi maka akan menurunkan kontribusi rumah tangga tersebut masuk dalam kategori yang memilih penyesuaian konsumsi makanan sebagai strategi *coping* sebesar 0,210.

b. Efek marginal variabel tindakan penyimpanan air (X_9)

$$\begin{aligned}
 \frac{\partial \hat{P}(y=0 | \mathbf{x})}{\partial X_9} &= -\beta_9 \phi(-\mathbf{x}'\boldsymbol{\beta}) \\
 &= -\beta_9 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\
 &= -0,512 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\
 &= -0,512 \phi(0,518) \\
 &= -0,512(0,349) \\
 &= -0,179 \\
 \\
 \frac{\partial \hat{P}(y=1 | \mathbf{x})}{\partial X_9} &= \beta_9 \phi(-\mathbf{x}'\boldsymbol{\beta}) \\
 &= \beta_9 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\
 &= 0,512 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\
 &= 0,512 \phi(0,518) \\
 &= 0,512(0,349) \\
 &= 0,179
 \end{aligned}$$

Nilai efek marginal variabel tindakan penyimpanan air sebesar -0,179 pada kategori $y=0$, hal ini menunjukkan bahwa pengaruh tindakan penyimpanan air terhadap keputusan rumah tangga di NTT dalam pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan adalah sebesar -0,179. Artinya rumah tangga yang melakukan penyimpanan air akan menurunkan kontribusi rumah tangga tersebut untuk masuk dalam kategori rumah tangga yang tidak memilih penyesuaian konsumsi makanan sebesar 0,179. Sedangkan pada $y=1$, variabel tindakan penyimpanan air akan meningkatkan kontribusi rumah

tangga tersebut untuk masuk dalam kategori yang memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan sebesar 0,179.

c. Efek marginal variabel pendapatan rumah tangga (X_{10})

$$\begin{aligned}\frac{\partial \hat{P}(y=0 | \mathbf{x})}{\partial X_{10}} &= -\beta_{10} \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= -\beta_{10} \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,815 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,815 \phi(0,518) \\ &= -0,815(0,349) \\ &= -0,284\end{aligned}$$

$$\begin{aligned}\frac{\partial \hat{P}(y=1 | \mathbf{x})}{\partial X_{10}} &= \beta_{10} \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= \beta_{10} \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,815 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,815 \phi(0,518) \\ &= 0,815(0,349) \\ &= 0,284\end{aligned}$$

Pada kategori $y=0$, nilai efek marginal variabel pendapatan rumah tangga kategori 1 sebesar -0,284. Hal ini menunjukkan bahwa pengaruh pendapatan rumah tangga dengan berbagai sumber pendapatan terhadap keputusan pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan adalah sebesar -0,284. Artinya rumah tangga yang memiliki berbagai sumber pendapatan akan menurunkan kontribusi rumah tangga tersebut untuk tidak memilih penyesuaian konsumsi makanan sebesar 0,284. Untuk $y=1$, diketahui rumah tangga yang memiliki berbagai sumber pendapatan akan menaikkan kontribusi rumah tangga tersebut

untuk memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan yaitu sebesar 0,284.

d. Efek marginal variabel sumber mata pencaharian utama bagi keluarga (X_{10})

Diperoleh nilai efek marginal variabel sumber mata pencaharian utama bagi keluarga sebesar -0,21, nilai tersebut menunjukkan bahwa pada $y=0$ pengaruh rumah tangga yang memiliki sumber pencaharian utama sebagai agro-pastoral terhadap keputusan pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan adalah sebesar -0,21. Artinya rumah tangga yang memiliki sumber mata pencaharian utama sebagai agro-pastoral akan menurunkan kontribusi rumah tangga tersebut untuk masuk dalam golongan rumah tangga yang tidak memilih penyesuaian konsumsi makanan sebesar 0,21.

$$\begin{aligned}\frac{\partial \hat{P}(y=0 | \mathbf{x})}{\partial X_{12}} &= -\beta_{12} \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= -\beta_{12} \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,601 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= -0,601 \phi(0,518) \\ &= -0,601(0,349) \\ &= -0,21\end{aligned}$$

$$\begin{aligned}\frac{\partial \hat{P}(y=1 | \mathbf{x})}{\partial X_{12}} &= \beta_{12} \phi(-\mathbf{x}'\boldsymbol{\beta}) \\ &= \beta_{12} \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,601 \phi(-0,807 - 0,603(1) + 0,512(1) + 0,815(1) + 0,601(1)) \\ &= 0,601 \phi(0,518) \\ &= 0,601(0,349) \\ &= 0,21\end{aligned}$$

Sedangkan pada $y=1$, rumah tangga yang memiliki sumber mata pencaharian utama sebagai agro-pastoral akan menaikkan kontribusi rumah tangga tersebut sebesar 0,21 untuk masuk dalam

golongan rumah tangga yang memilih penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan.

4.4.5 Ketepatan Klasifikasi

Hasil ketepatan klasifikasi pada data faktor-faktor yang mempengaruhi dalam pemilihan keputusan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko di NTT adalah ditunjukkan pada Tabel 4.6. Tabel 4.6 menunjukkan bahwa model yang terbentuk beberapa belum tepat diklasifikasikan antara observasi dengan hasil prediksi. Sebanyak 148 rumah tangga tepat diklasifikasikan memilih untuk melakukan penyesuaian konsumsi makanan sebagai strategi *coping* manajemen risiko kekeringan di NTT dan 51 rumah tangga tepat diklasifikasikan tidak memilih strategi *coping* tersebut dari total 300 rumah tangga.

Tabel 4.6 Ketepatan Klasifikasi

	Observasi	Prediksi		Total
		Tidak	Ya	
Pemilihan penyesuaian konsumsi makanan sebagai strategi <i>coping</i>	Tidak	51	58	109
	Ya	43	148	191
	Total	94	206	300

Berdasarkan hasil ketepatan klasifikasi pada Tabel 4.10, maka dapat dihitung persentase ketepatan klasifikasi model sebagai berikut.

$$\text{Akurasi} = \frac{\sum (\text{prediksi benar})}{\sum (\text{semua prediksi})} = \frac{51 + 148}{300} = 0,663$$

Hasil akurasi diperoleh senilai 0,663 atau sebesar 66,3%. Nilai akurasi tersebut menunjukkan bahwa dengan model regresi probit mampu mengklasifikasikan data pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko

kekeringan di NTT dengan hasil prediksi tepat adalah sebesar 66,3%.

4.5 Random Forest

Pada penelitian ini untuk analisis *random forest* terdapat 300 data yang terbagi menjadi data latih (*train*) sebanyak 225 data dan sisanya sebanyak 75 data sebagai data uji (*test*). Berikut adalah hasil dari analisis *random forest* yang diperoleh.

4.5.1 Penentuan *Mtry* dan *Ntree* Optimum

Pada metode *random forest* setelah membagi data *training* dan data *testing*, langkah selanjutnya yaitu menentukan nilai *mtry* (peubah penjelas) dan *nree* (jumlah pohon) terbaik agar model yang didapat optimal dengan melihat nilai *error* OOB yang paling kecil. Penentuan *mtry* dapat dilakukan dengan tiga cara seperti persamaan (2.30), (2.31) dan (2.32) dengan hasil sebagai berikut.

$$Mtry = \frac{1}{2} \left| \sqrt{p} \right| = \frac{1}{2} \left| \sqrt{12} \right| = 1,73 = 2$$

$$Mtry = \left| \sqrt{p} \right| = \left| \sqrt{12} \right| = 3,46 = 3$$

$$Mtry = 2 \times \left| \sqrt{p} \right| = 2 \times \left| \sqrt{12} \right| = 6,92 = 7$$

Setelah menentukan nilai *mtry*, maka dapat dilanjutkan dengan mencoba *mtry* yang masing-masing telah didapatkan menggunakan *default* jumlah pohon (*nree*) yaitu 500, sehingga didapatkan nilai *error* OOB sebagai berikut :

Tabel 4.7 Pengujian *Mtry* Terbaik

<i>Mtry</i>	<i>Error</i> OOB
2	32,89%
3	33,78%
7	32%

Tabel 4.10 diatas menunjukkan nilai *error* OOB terkecil yaitu 32% didapatkan pada saat *mtry* sebesar 7. Sehingga *mtry* optimum yang digunakan adalah sebesar 7, yang selanjutnya dapat digunakan untuk mendapatkan jumlah pohon (*nree*) optimum dengan mencari nilai *error* OOB terkecil. Jumlah pohon

(*ntree*) yang akan dicoba dimulai dari 25, 50, 250, 500, 1000, 1500 dan 2000 dengan menggunakan nilai *mtry* sebesar 7. Hasil yang diperoleh adalah sebagai berikut.

Tabel 4.8 Pengujian Nilai OOB dengan *Ntree* Berbeda

<i>Ntree</i>	<i>Error OOB</i>
25	36,44%
50	35,11%
250	33,78%
500	35,11%
1000	32,89%
1500	35,11%
2000	35,56%

Tabel 4.8 menunjukkan dari 5 *ntree* yang dicobakan, nilai *ntree* optimum atau terbaik adalah sebesar 1000 dengan menghasilkan nilai *error* terkecil yaitu 32,89%. Setelah mendapatkan nilai *mtry* dan *ntree* optimum, maka nilai tersebut dapat digunakan untuk menentukan prediksi *random forest* untuk data *training*.

4.5.2 Pembentukan Model

Dalam pembentukan model *random forest* terbaik, hasil iterasi harus optimal karena hasil iterasi adalah jumlah pohon yang akan digunakan dalam *learning*. Pada Lampiran 19 menunjukkan bahwa dengan tipe *random forest* yang digunakan adalah klasifikasi, *mtry* yang digunakan sebesar 7 dan *ntree* sebesar 1000 yang diambil dari hasil pemilihan *mtry* serta *ntree* optimum sebelumnya didapatkan nilai estimasi *error OOB* dengan menggunakan data *training* adalah sebesar 32,89%.

Tabel 4.9 Hasil Prediksi Data Latih (*Train*)

		Memilih Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan		<i>Class Error</i>
		Tidak	Ya	
Memilih Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan	Tidak	47	42	0,4719
	Ya	32	104	0,2353

Berdasarkan Tabel 4.9, menunjukkan bahwa banyaknya rumah tangga yang tepat diprediksi memilih strategi *coping* penyesuaian konsumsi makanan adalah sebanyak 104 rumah tangga dan tepat diprediksi tidak memilih strategi *coping* penyesuaian konsumsi makanan adalah sebanyak 47 rumah tangga.

$$\text{Akurasi} = \frac{\sum(\text{prediksi benar})}{\sum(\text{semua prediksi})} = \frac{47+104}{225} = 0,671$$

Hasil akurasi untuk data latih (*train*) diperoleh senilai 0,671 atau sebesar 67,1%. Nilai akurasi tersebut menunjukkan bahwa random forest melakukan prediksi berdasarkan data latih (*train*) dengan hasil prediksi tepat adalah sebesar 67,1%.

4.5.3 Perhitungan Akurasi

Setelah membuat model pada *random forest*, langkah selanjutnya yaitu melakukan perhitungan akurasi untuk data uji (*test*) dengan hasil sebagai berikut.

Tabel 4.10 Hasil Prediksi Data Uji (*Test*)

		Memilih Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan		<i>Class Error</i>
		Tidak	Ya	
Memilih Strategi <i>Coping</i> Penyesuaian Konsumsi Makanan	Tidak	10	10	0,5000
	Ya	8	47	0,1454

Tabel 4.10 diatas dengan menggunakan data uji (*test*) menunjukkan bahwa banyaknya rumah tangga yang tepat diprediksi memilih strategi *coping* penyesuaian konsumsi makanan adalah sebanyak 47 rumah tangga dan tepat diprediksi

tidak memilih strategi *coping* penyesuaian konsumsi makanan adalah sebanyak 10 rumah tangga.

$$\text{Akurasi} = \frac{\sum (\text{prediksi benar})}{\sum (\text{semua prediksi})} = \frac{10+47}{75} = 0,76$$

Hasil akurasi untuk data uji (*test*) menunjukkan nilai yang lebih tinggi dibandingkan nilai akurasi berdasarkan data latih (*train*) yaitu diperoleh nilai 0,76 atau sebesar 76%. Nilai akurasi tersebut menunjukkan bahwa *random forest* melakukan prediksi berdasarkan data uji (*test*) dengan hasil prediksi tepat adalah sebesar 76% yang dapat dikatakan cukup bagus.

4.5.4 Importance Variable

Pada analisis *random forest* diketahui nilai akurasi terbesar didapatkan dengan menggunakan data uji (*test*), maka dalam menentukan variabel kepentingan (*importance variabel*) dilakukan dengan menggunakan data uji (*test*).

Tabel 4.11 Importance Variable

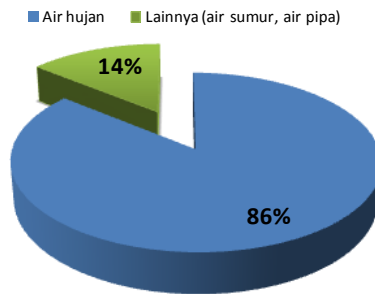
Variabel	Mean Decrease Gini
Kepemilikan lahan pribadi (x_4)	7,348
Usia kepala rumah tangga (x_2)	4,346
Total aset yang dimiliki (x_3)	4,097
Total hewan yang dimiliki lima tahun terakhir (x_6)	3,938
Sumber mata pencaharian utama bagi keluarga (x_{11})	3,222
Proporsi anggota keluarga yang tidak aktif dalam ekonomi (x_5)	2,544
Kepemilikan akses tanah untuk bercocok tanam atau menggembala ternak (x_{12})	1,328
Pendapatan rumah tangga (x_{10})	0,688
Tindakan penyimpanan air (x_9)	0,518
Respon terhadap ramalan cuaca (x_8)	0,454
Manajemen ternak (x_7)	0,344
Jenis kelamin kepala rumah tangga (x_1)	0,071

Tabel 4.11 diatas menunjukkan hasil *importance variable* yang terbentuk dari hasil klasifikasi *random forest*. *Importance variable* menunjukkan hubungan variabel itu sendiri dalam mempengaruhi hasil analisis atau prediksi yang dilakukan, dimana semakin besar nilainya maka semakin besar peranan variabel tersebut dalam mempengaruhi hasil analisis. Tabel 4.14 menunjukkan variabel yang memiliki peranan paling penting adalah kepemilikan lahan pribadi (x_4), usia kepala rumah tangga (x_2), lalu diikuti oleh total aset yang dimiliki (x_3), total hewan yang dimiliki lima tahun terakhir (x_6), sumber mata pencaharian utama bagi keluarga (x_{11}), hingga yang paling kecil peranannya adalah jenis kelamin kepala rumah tangga (x_1).

4.6 Penerapan Kasus

Data survei kasus kekeringan di Nusa Tenggara Timur menunjukkan bahwa kekeringan sangat mempengaruhi hasil perekonomian karena berdampak langsung pada produktivitas pertanian maupun peternakan. Saat kekeringan terjadi, lebih dari 90% rumah tangga di NTT merasakan pengaruh yang besar yang disebabkan oleh kekeringan yaitu antara lain mulai dari produktivitas pertanian menurun, sulit mendapatkan makanan, penghasilan merosot hingga terjadi kelaparan. Untuk dapat bertahan selama terjadi kekeringan dan mencegah dampak negatif akibat kekeringan seperti tahun-tahun sebelumnya, rumah tangga di Nusa Tenggara Timur dapat menerapkan strategi *coping* salah satunya yaitu melakukan penyesuaian konsumsi makanan. Bagi rumah tangga yang menerapkan strategi *coping* penyesuaian konsumsi makanan dapat meminimalkan dampak negatif saat terjadi kekeringan, penerapan strategi tersebut dapat dijadikan langkah antisipasi untuk menghindari kesulitan makanan maupun kelaparan. Saat terjadi kekeringan, Rumah tangga yang menerapkan strategi *coping* tersebut dapat tetap makan meskipun proporsi atau kualitas makanan yang dikonsumsi berbeda dengan yang dikonsumsi saat tidak terjadi kekeringan sehingga resiko kelaparan juga akan semakin menurun. Penyesuaian konsumsi makanan sebaiknya hanya dilakukan dalam jangka pendek saja

yaitu saat kekeringan berlangsung, karena apabila dilakukan dalam jangka panjang akan dikhawatirkan menyebabkan malnutrisi parah dan meningkatkan risiko kematian bayi menurut Schilderink pada tahun 2009. Pemerintah dapat melakukan pengawasan serta sosialisasi terhadap rumah tangga yang melakukan penyesuaian konsumsi makanan sebagai strategi coping terhadap risiko kekeringan dengan kriteria rumah tangga tidak memiliki lahan pribadi, memiliki penyimpanan air, memiliki berbagai sumber pendapatan, dan rumah tangga bermata pencaharian pada bidang agro-pastoral untuk lebih memperhatikan tindakan strategi *coping* mereka agar tepat pada jangka waktu yang seharusnya.



Gambar 4.10 Sumber Air untuk Tanaman

Berdasarkan hasil analisis yang didapatkan salah satu hal yang dapat dilakukan oleh Pemerintah NTT untuk kedepannya adalah memperhatikan mengenai penyimpanan air yang dimiliki oleh setiap rumah tangga. Pada Gambar 4.6 menunjukkan bahwa hasil survei kasus kekeringan di NTT tahun 2018 rumah tangga yang tidak memiliki penyimpanan cukup banyak yaitu sebesar 40,3%. Dari 40,3% rumah tangga tersebut ternyata mayoritas rumah tangganya masih mengandalkan air hujan untuk mengaliri tanaman (pertanian) mereka, padahal jumlah bulan hujan di Nusa Tenggara Timur hanya selama 4 bulan dan itupun tidak setiap hari turun hujan. Sehingga rumah tangga yang mengandalkan air hujan untuk mengaliri tanaman mereka akan sangat kesulitan air

pada bulan-bulan masa tanam diluar bulan musim hujan. Jika air yang dibutuhkan untuk pengairan lahan pertanian tidak mencukupi, maka produktivitas panen petani akan menurun hingga kemungkinan gagal panen. Ketika stok pangan mulai menipis harga untuk kebutuhan pokok di pasaran akan menyebabkan harga melonjak dari harga normal, karena persediaan yang lebih kecil dari permintaan. Sehingga bukan tidak mungkin bahwa apabila harga kebutuhan pokok menjadi mahal akan menyebabkan kelaparan di rumah tangga lapisan bawah karena tidak mampu membeli makanan dan kebutuhan lainnya. Maka oleh sebab itulah, pemerintah setempat dapat menghimbau rumah tangga di NTT akan pentingnya tindakan penyimpanan air saat musim hujan dan pemerintah juga dapat membantu dalam membangun penyimpanan air untuk penduduknya seperti memperbanyak membangun embung (penyimpanan air hujan), membuat pipa air yang lebih banyak, dan melakukan pengerukan bendungan yang sudah dangkal. Kepala Dinas Pekerjaan Umum Provinsi Nusa Tenggara Timur Andre Koreh mengatakan bahwa pada tahun 2018 baru ada 1.125 embung, padahal jumlah embung yang dibutuhkan adalah sekitar 4.000 embung agar dapat mengatasi kebutuhan air di Nusa Tenggara Timur. Langkah tersebut dapat membantu irigasi lahan pertanian dan dapat membantu ketersediaan air rumah tangga di NTT agar dapat bertahan saat terjadi kekeringan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang didapatkan pada analisis faktor-faktor yang mempengaruhi pemilihan strategi penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan di NTT menunjukkan mayoritas sebanyak 64% (191 rumah tangga) memilih melakukan penyesuaian konsumsi makanan sebagai strategi *coping* dalam mengurangi risiko kekeringan pada tahun 2018, dengan didominasi oleh kepala rumah tangga laki-laki yaitu 174 kepala rumah tangga yang rata-rata berusia 49 tahun. Dari 191 rumah tangga di NTT yang melakukan penyesuaian konsumsi makanan sebagai strategi *coping*, 170 rumah tangga memiliki lahan pribadi, 165 melakukan manajemen ternak saat kekeringan, tetapi hanya minoritas yang sudah merespon ramalan cuaca yaitu sebanyak 27 rumah tangga, selain itu terdapat 113 rumah tangga belum memiliki akses tanah untuk bercocok tanam maupun menggembala ternak. Dari segi pendapatan dan pekerjaan, mayoritas sumber pendapatan rumah tangga adalah tunggal dengan mata pencaharian utama dari agro-pastoral. Hasil regresi probit menunjukkan rumah tangga dalam memilih keputusan untuk melakukan penyesuaian konsumsi makanan sebagai strategi *coping* dipengaruhi oleh kepemilikan lahan pribadi, pendapatan rumah tangga, sumber mata pencaharian utama bagi keluarga dan kepemilikan akses tanah dengan ketepatan klasifikasi sebesar 66,3%. Pada analisis *random forest* menunjukkan tingkat klasifikasi sebesar yaitu 76%, dan variabel yang memiliki peranan paling penting adalah kepemilikan lahan pribadi.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selanjutnya yaitu metode *random forest* dapat digunakan untuk klasifikasi dalam analisis kasus pemilihan strategi *coping* terhadap risiko kekeringan karena pada penelitian ini metode tersebut

menghasilkan nilai akurasi paling besar untuk hasil klasifikasi. Selain itu, bagi pemerintah Nusa Tenggara Timur dapat menggunakan hasil analisis regresi probit biner sebagai bahan pertimbangan pembuatan kebijakan pada kasus pemilihan penyesuaian konsumsi makanan sebagai strategi *coping* terhadap risiko kekeringan khususnya.

DAFTAR PUSTAKA

- Aldrian, E. (2008). *Meteorologi Laut Indonesia*. Jakarta: Badan Meteorologi dan Geofisika.
- Agresti, A. (2002). *Categorical Data Analysis* (2 ed.). New York: John Wiley & Sons, Inc.
- As-syakur, A. R., Adnyana, I., Mahendra, M., Arthana, I., Merit, I., Kasa, I., et al. (2013). Observations of spatial patterns on the rainfall response to ENSO and IOD over Indonesia using TRMM Multisatellite Precipitation Analysis (TMPA). *International Journal of Climatology*, pp. 1-13.
- Barros, R. C., de Carvalho, A. C., & Freitas, A. (2015). In Automatic Design of Decision-Tree Induction Algorithms. *Decision-Tree Induction* , pp. 7-45.
- BKP. (2014). Laporan Kinerja Badan Ketahanan Pangan Tahun 2010-2014. Jakarta: Badan Ketahanan Pangan, Kementerian Pertanian.
- BNPB. (2018). *Data Informasi Bencana Indonesia*. Diakses pada tanggal 20 Februari 2019. <http://bnpb.cloud/dibi/tabel1b>.
- Breiman, L. (2001). Statistical modeling : The two cultures (with comments and a rejoinder by author). *Statistical science*, Volume 16(3), pp. 199-231.
- Breiman, L., & Cutler, A. (2003). Random Forests. *Machine Learning*. Volume 40, pp. 5-32.
- Greene, W. H. (2008). *Econometrics Analysis* (6 ed.). New Jersey: Prentice Hall, Inc.
- Gujarati, D. N. (2003). *Basic Econometric* (4 ed.). New York: Mc Graw-Hill.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Technique*. Waltham: Elsevier Inc.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2 ed.). New York: John Willey & Sons, Inc.

- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6 ed.). New Jersey: Prentice Hall.
- Kertamuda, F., & H., H. (2009). Pengaruh Strategi Coping Terhadap Penyesuain Dirinya Mahasiswa Baru. *Jurnal Universitas Paradina*, pp. 14.
- Kotsiantis, S. (2013). Decision Tree: A Recent Overview. *Artificial Intelligence Review* , pp. 261-283.
- Liverman, D. M. (1999). Vulnerability and Adaptation to Drought in Mexico. *39 Nat Resources J.* , pp. 109-114.
- Longford, N. T. (2005). *Missing Data and Small-Area Estimation*. New York: Springer.
- Mavhura, E., Manatsa, D., & Mushore, T. (2015). Adaptation to drought in arid and semi-arid environments: Case of the Zambezi Valley, Zimbabwe. *Jamba: Journal of Disaster Risk Studies*, Volume 7(1), pp. 144.
- Mishra, S. (2007). Household Livelihood and Coping Mechanism During Drought. *J. Soc. Sci.*, Volume 15(2), pp. 181-186.
- Nagarajan, R. (2009). *Drought Assessment*. Dordrecht-Netherlands: Springer.
- Nagler, J. (1994). Interpreting Probit Analysis. Diakses pada tanggal 20 Februari 2019. <http://www.nyu.edu/classes/nagler/quant2/notes/probit1.pdf>
- Pandey, S., & Bhandari, H. (2009). Drought, coping mechanisms and poverty: Insights from rainfed rice farming in Asia. *IFAD Discussion Paper Series 7*, pp. 6-42.
- Schilderink, G. (2009). Drought Cycle Management in Arid and Semi-Arid Kenya : A Relevant Disaster risk reduction model ?. *Development Economist*, pp. 6-49.
- Schouten, K., Frasinca, F., & Dekker, R. (2016). An Information gain-Driver Feature Study for Aspect-Based Sentiment Anaysisi. *Natural Language Processing and Information Systems*, pp. 2-12.

- Stuart, G. W. & Sundeen, S. J., (1991). *Pocket guide to psychiatric nursing*. 3 ed. Toronto: The Mosby Company.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi ke-3*. Terjemahan: Bambang Sumantri. Jakarta: PT. Gramedia Pustaka Utama.
- Winarno, W. W. (2007). *Analisis Ekonometrika dan Statistik dengan EVIEWS*. Yogyakarta: UPP STIM YKPN.
- Thywissen. (2006). *Components of Risk, A Comparative Glosary*. Bonn, Germany: UNU Institute for Environment and Human Security (UNU-EHS).
- UNISDR. (2011). Menguak Risiko, Menggagas Makna Baru Pembangunan. *Laporan Pengkajian Global Tentang Pengurangan Risiko Bencana*, pp. 5.

Halaman ini sengaja dikosongkan

LAMPIRAN

Lampiran 1. Surat Pernyataan Pengambilan Data

SURAT PERNYATAAN

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

Nama : Pratiwi Penta Atrivi

NRP : 0621174500009

menyatakan bahwa data yang digunakan dalam Tugas Akhir ini merupakan data sekunder yang diambil dari hasil survei “*Drought Cycle Management*” Pusat Studi Kebumihan, Bencana, dan Perubahan Iklim ITS oleh Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.

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

Mengetahui
Pembimbing Tugas Akhir



Dr. rer. pol. Heri Kuswanto, S.Si., M.Si
NIP. 19820326 200312 1 004

Surabaya, 27 Mei 2019



Pratiwi Penta Atrivi
NRP.0621174500009

Lampiran 2. Data Faktor-Faktor yang Mempengaruhi Pemilihan Keputusan Penyesuaian konsumsi makanan rumah Tangga Sebagai Strategi *Coping* Terhadap Risiko Kekeringan di NTT

No	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	Y
1	0	42	3	1	0.4	10	1	0	1	0	1	0	1
2	0	54	2	1	0.5	25	1	0	1	0	1	0	1
3	0	47	2	1	0.429	10	1	0	1	1	1	0	0
4	0	52	2	1	0.5	10	1	0	1	1	1	0	0
5	0	58	2	1	0.429	7	1	0	1	1	1	0	0
6	0	65	2	1	0	8	1	0	1	0	1	0	0
7	0	66	2	1	0	23	1	0	1	0	1	0	1
8	0	58	2	1	0	15	1	0	1	0	1	0	1
9	0	60	2	0	0.4	13	1	0	1	0	1	0	1
10	0	40	2	0	0.5	22	1	0	1	0	1	0	1
11	0	50	2	0	0	0	0	0	0	1	2	0	1
12	0	39	6	1	0.333	30	1	0	0	1	1	0	0
13	0	37	3	1	0.25	5	1	0	1	1	1	0	1
14	0	53	2	0	0.333	30	1	0	0	0	1	0	1
15	0	38	5	1	0	5	1	1	1	1	1	0	1
16	0	52	2	1	0	10	1	0	1	0	1	0	1
17	0	50	2	1	0	65	1	0	1	0	1	0	1
18	0	42	3	1	0.6	10	1	0	1	0	1	0	1
19	0	48	3	1	0.667	10	1	0	1	0	1	0	1
20	0	30	2	1	0.333	20	1	0	1	0	1	0	1
21	0	37	3	1	0.333	1	1	1	1	1	1	0	1
22	0	53	2	1	0.5	0	0	0	1	0	1	0	1
23	0	64	2	1	0	10	1	0	1	0	1	0	1
24	0	50	3	1	0.6	45	1	0	1	0	1	0	1
25	0	48	2	1	0.6	10	1	0	1	0	1	0	1
26	0	50	3	1	0	5	1	0	1	0	1	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
297	0	59	5	1	0	22	1	1	1	0	0	1	1
298	0	67	4	0	0	12	1	0	0	0	1	1	0
299	0	62	4	1	0	3	1	0	1	0	1	1	1
300	0	62	5	0	0	2	1	0	0	0	1	1	0

Lampiran 3. Pengecekan dan Penanganan *Missing Value*

Variable	Total Count	N*
x1	300	0
x2	300	2
x3	300	0
x4	300	0
x5	300	0
x6	300	0
x7	300	0
x8	300	0
x9	300	1
x10	300	0
x11	300	0
x12	300	0
y	300	0

Variable	Total Count	Mean
x2	300	48.748

Variable	Total Count	Mode	N for Mode
x9	300	1	178

Variable	Total Count	N*
x1	300	0
x2	300	0
x3	300	0
x4	300	0
x5	300	0
x6	300	0
x7	300	0
x8	300	0
x9	300	0
x10	300	0
x11	300	0
x12	300	0
y	300	0

Lampiran 4. Statistika Deskriptif Data Usia Kepala Rumah Tangga (X_2), Total Aset yang Dimiliki (X_3), Proporsi Anggota Keluarga yang Tidak Aktif Dalam Ekonomi (X_5), dan Total Hewan yang Dimiliki Lima Tahun Terakhir (X_6)

Variable	Total					
	Count	Mean	StDev	Minimum	Median	Maximum
x2	300	48.748	11.803	25.000	48.875	88.000
x3	300	3.8100	1.5082	0.0000	4.0000	10.0000
x5	300	0.3977	0.3695	0.0000	0.3330	3.0000
x6	300	20.68	26.34	0.00	11.00	220.00

Lampiran 5. Tabel Frekuensi Pemilihan Strategi *Coping* Penyesuaian Konsumsi Makanan (Y)

Y

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tidak	109	36.3	36.3	36.3
	Ya	191	63.7	63.7	100.0
	Total	300	100.0	100.0	

Lampiran 6. *Crosstab* Jenis Kelamin Kepala Rumah Tangga Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X1	Laki-laki	Count	103	174	277
		Expected Count	100.6	176.4	277.0
	Perempuan	Count	6	17	23
		Expected Count	8.4	14.6	23.0
Total		Count	109	191	300
		Expected Count	109.0	191.0	300.0

Lampiran 7. *Crosstab* Kepemilikan Lahan Pribadi Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X4	Tidak	Count	24	21	45
		Expected Count	16.4	28.7	45.0
	Ya	Count	85	170	255
		Expected Count	92.7	162.4	255.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 8. *Crosstab* Manajemen Ternak Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X7	Tidak	Count	16	26	42
		Expected Count	15.3	26.7	42.0
	Ya	Count	93	165	258
		Expected Count	93.7	164.3	258.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 9. *Crosstab* Respon Terhadap Ramalan Cuaca Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X8	Tidak	Count	85	164	249
		Expected Count	90.5	158.5	249.0
	Ya	Count	24	27	51
		Expected Count	18.5	32.5	51.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 10. *Crosstab* Tindakan Penyimpanan Air Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X9	Tidak	Count	42	79	121
		Expected Count	44.0	77.0	121.0
	Ya	Count	67	112	179
		Expected Count	65.0	114.0	179.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 11. *Crosstab* Pendapatan Rumah Tangga Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X10	Sumber Pendapatan Tunggal	Count	46	136	182
		Expected Count	66.1	115.9	182.0
	Berbagai Sumber Pendapatan	Count	63	55	118
		Expected Count	42.9	75.1	118.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 12. *Crosstab* Sumber Mata Pencarian Utama Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X11	Pastoral	Count	16	12	28
		Expected Count	10.2	17.8	28.0
	Agro-Pastoral	Count	80	166	246
		Expected Count	89.4	156.6	246.0
	Bisnis Skala Kecil	Count	6	2	8
		Expected Count	2.9	5.1	8.0
	Upah Kerja	Count	7	11	18
		Expected Count	6.5	11.5	18.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 13. *Crosstab* Kepemilikan Akses Tanah Untuk Bercocok Tanam atau Menggembala Ternak Berdasarkan Pemilihan Strategi *Coping*

			Y		Total
			Tidak	Ya	
X12	Tidak	Count	41	113	154
		Expected Count	56.0	98.0	154.0
	Ya	Count	68	78	146
		Expected Count	53.0	93.0	146.0
Total	Count		109	191	300
	Expected Count		109.0	191.0	300.0

Lampiran 14. *Output Minitab* Metode Regresi Probit Model Awal

Variable	Value	Count	
y	0	109	(Event)
	1	191	
	Total	300	

Deviance	Deviance	
R-Sq	R-Sq(adj)	AIC
15.59%	12.03%	361.88

Lampiran 14. *Output Minitab* Metode Regresi Probit Model Awal (lanjutan)

Deviance Table							
Source	DF	Seq Dev	Contribution	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	14	61.306	15.59%	61.306	4.3790	61.31	0.000
x2	1	1.244	0.32%	0.016	0.0161	0.02	0.899
x3	1	0.383	0.10%	2.033	2.0332	2.03	0.154
x5	1	3.436	0.87%	1.286	1.2863	1.29	0.257
x6	1	1.849	0.47%	0.154	0.1539	0.15	0.695
x1	1	0.923	0.23%	0.379	0.3788	0.38	0.538
x4	1	6.501	1.65%	5.781	5.7811	5.78	0.016
x7	1	0.059	0.02%	0.006	0.0063	0.01	0.937
x8	1	2.701	0.69%	0.801	0.8011	0.80	0.371
x9	1	0.474	0.12%	7.689	7.6885	7.69	0.006
x10	1	22.440	5.71%	14.393	14.3928	14.39	0.000
x11	3	3.716	0.94%	8.133	2.7110	8.13	0.043
x12	1	17.581	4.47%	17.581	17.5809	17.58	0.000
Error	285	331.881	84.41%	331.881	1.1645		
Total	299	393.187	100.00%				
Coefficients							
Term	Coef	SE Coef	95% CI	Z-Value	P-Value	VIF	
Constant	-0.444	0.621	(-1.661, 0.773)	-0.71	0.475		
x2	-0.00090	0.00710	(-0.01481, 0.01301)	-0.13	0.899	1.13	
x3	-0.0843	0.0593	(-0.2006, 0.0319)	-1.42	0.155	1.27	
x5	0.253	0.226	(-0.189, 0.696)	1.12	0.262	1.11	
x6	0.00139	0.00352	(-0.00552, 0.00830)	0.39	0.693	1.27	
x1							
1	-0.206	0.338	(-0.869, 0.457)	-0.61	0.542	1.04	
x4							
1	-0.546	0.228	(-0.993, -0.100)	-2.40	0.016	1.13	
x7							
1	0.020	0.245	(-0.461, 0.500)	0.08	0.937	1.16	
x8							
1	0.194	0.216	(-0.230, 0.617)	0.90	0.370	1.11	
x9							
1	0.538	0.197	(0.152, 0.924)	2.73	0.006	1.48	
x10							
1	0.682	0.181	(0.327, 1.037)	3.76	0.000	1.29	
x11							
1	-0.320	0.294	(-0.896, 0.256)	-1.09	0.276	2.10	
2	1.012	0.581	(-0.126, 2.150)	1.74	0.081	1.51	
3	-0.324	0.434	(-1.174, 0.527)	-0.75	0.456	1.72	
x12							
1	0.747	0.181	(0.392, 1.102)	4.12	0.000	1.29	

Lampiran 14. Output Minitab Metode Regresi Probit Model Awal (lanjutan)

Regression Equation

$$P(0) = \Phi(Y')$$

$$Y' = -0.444 - 0.00090 x_2 - 0.0843 x_3 + 0.253 x_5 + 0.00139 x_6 + 0.0 x_{1_0} - 0.206 x_{1_1} + 0.0 x_{4_0} - 0.546 x_{4_1} + 0.0 x_{7_0} + 0.020 x_{7_1} + 0.0 x_{8_0} + 0.194 x_{8_1} + 0.0 x_{9_0} + 0.538 x_{9_1} + 0.0 x_{10_0} + 0.682 x_{10_1} + 0.0 x_{11_0} - 0.320 x_{11_1} + 1.012 x_{11_2} - 0.324 x_{11_3} + 0.0 x_{12_0} + 0.747 x_{12_1}$$

Φ = CDF of the standard normal distribution

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	285	331.88	0.029
Pearson	285	316.79	0.095
Hosmer-Lemeshow	8	12.91	0.115

Measures of Association

Pairs	Number	Percent	Summary Measures	Value
Concordant	15952	76.6	Somers' D	0.54
Discordant	4797	23.0	Goodman-Kruskal Gamma	0.54
Ties	70	0.3	Kendall's Tau-a	0.25
Total	20819	100.0		

Association is between the response variable and predicted probabilities

Lampiran 15. Output Minitab Metode Regresi Probit Model Akhir

Binary Logistic Regression: y versus x4, x9, x10, x12

Method

Link function	Normit
Categorical predictor coding	(1, 0)
Rows used	300

Lampiran 15. Output Minitab Metode Regresi Probit Model Akhir (lanjutan)

Response Information							
Variable	Value	Count					
y	0	109	(Event)				
	1	191					
	Total	300					

Deviance Table							
Source	DF	Seq Dev	Contribution	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	4	48.715	12.39%	48.715	12.179	48.71	0.000
x4	1	6.382	1.62%	8.076	8.076	8.08	0.004
x9	1	0.333	0.08%	8.714	8.714	8.71	0.003
x10	1	28.877	7.34%	25.380	25.380	25.38	0.000
x12	1	13.123	3.34%	13.123	13.123	13.12	0.000
Error	295	344.472	87.61%	344.472	1.168		
Total	299	393.187	100.00%				

Model Summary			
Deviance	Deviance		
R-Sq	R-Sq(adj)	AIC	
12.39%	11.37%	354.47	

Coefficients						
Term	Coef	SE Coef	95% CI	Z-Value	P-Value	VIF
Constant	-0.807	0.258	(-1.312, -0.302)	-3.13	0.002	
x4						
1	-0.603	0.213	(-1.021, -0.186)	-2.83	0.005	1.01
x9						
1	0.512	0.176	(0.167, 0.856)	2.91	0.004	1.21
x10						
1	0.815	0.164	(0.493, 1.136)	4.97	0.000	1.09
x12						
1	0.601	0.168	(0.273, 0.930)	3.59	0.000	1.14

Regression Equation	
$P(0)$	$= \Phi(Y')$
Y'	$= -0.807 + 0.0 x4_0 - 0.603 x4_1 + 0.0 x9_0 + 0.512 x9_1 + 0.0 x10_0 + 0.815 x10_1$ $+ 0.0 x12_0 + 0.601 x12_1$
Φ	$=$ CDF of the standard normal distribution

Lampiran 15. Output Minitab Metode Regresi Probit Model Akhir (lanjutan)

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	295	344.47	0.025
Pearson	295	309.26	0.272
Hosmer-Lemeshow	5	11.33	0.045

Measures of Association

Pairs	Number	Percent	Summary Measures	Value
Concordant	14112	67.8	Somers' D	0.46
Discordant	4527	21.7	Goodman-Kruskal Gamma	0.51
Ties	2180	10.5	Kendall's Tau-a	0.21
Total	20819	100.0		

Association is between the response variable and predicted probabilities

Lampiran 16. Probabilitas Pemilihan Strategi *Coping* Penyesuaian Konsumsi Makanan

Rumah Tangga	Y=0 (Tidak Memilih)	Y=1 (Memilih)	Prediksi	Aktual
1	0.184429198	0.815570802	1	1
2	0.184429198	0.815570802	1	1
3	0.466506349	0.533493651	1	0
4	0.466506349	0.533493651	1	0
5	0.466506349	0.533493651	1	0
6	0.184429198	0.815570802	1	0
7	0.184429198	0.815570802	1	1
8	0.184429198	0.815570802	1	1
9	0.383855425	0.616144575	1	1
10	0.383855425	0.616144575	1	1
11	0.503060182	0.496939818	0	1
12	0.275730431	0.724269569	1	0
13	0.466506349	0.533493651	1	1
14	0.209865507	0.790134493	1	1
15	0.466506349	0.533493651	1	1
16	0.184429198	0.815570802	1	1
17	0.184429198	0.815570802	1	1
18	0.184429198	0.815570802	1	1
19	0.184429198	0.815570802	1	1
20	0.184429198	0.815570802	1	1
21	0.466506349	0.533493651	1	1
22	0.184429198	0.815570802	1	1
23	0.184429198	0.815570802	1	1
24	0.184429198	0.815570802	1	1
25	0.184429198	0.815570802	1	1
26	0.184429198	0.815570802	1	1
27	0.184429198	0.815570802	1	1
⋮	⋮	⋮	⋮	⋮
297	0.383180417	0.616819583	1	1
298	0.418624889	0.581375111	1	0
299	0.383180417	0.616819583	1	1
300	0.418624889	0.581375111	1	0

Lampiran 17. Tabulasi Silang Klasifikasi Antara Observasi dan Hasil Prediksi Model Regresi Probit Biner

Rows: Aktual		Columns: Prediksi		
	0	1	All	
0	51 17.00 34.15	58 19.33 74.85	109 36.33	
1	43 14.33 59.85	148 49.33 131.15	191 63.67	
All	94 31.33	206 68.67	300 100.00	
<i>Cell Contents</i>				
<i>Count</i>				
<i>% of Total</i>				
<i>Expected count</i>				

Lampiran 18. *Syntax Random Forest*

```
library(randomForest)

#memanggil data
mydata <- read.csv("E:/datakekeringan.csv")
attach(mydata)

#mengkonversikan variabel x1, x4, x7, x8, x9, x10, x11, x12, dan y
kedalam bentuk variabel kategorik
mydata$x1 <- factor(mydata$x1)
mydata$x4 <- factor(mydata$x4)
```

Lampiran 18. Syntax Random Forest (lanjutan)

```

mydata$x7 <- factor(mydata$x7)
mydata$x8 <- factor(mydata$x8)
mydata$x9 <- factor(mydata$x9)
mydata$x10 <- factor(mydata$x10)
mydata$x11 <- factor(mydata$x11)
mydata$x12 <- factor(mydata$x12)
mydata$y <- factor(mydata$y)

#training testing
set.seed(123)
train <- mydata[1:225, ]
test <- mydata[226:300, ]

#mencari mtry optimum dengan menggunakan ntree default=500
coba.mtry2 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=2, ntree=500,
importance=TRUE, na.action=na.omit)
coba.mtry3 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=3, ntree=500,
importance=TRUE, na.action=na.omit)
coba.mtry7 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=500,
importance=TRUE, na.action=na.omit)
coba.mtry2
coba.mtry3
coba.mtry7

#mencari ntree optimum dengan menggunakan mtry optimum=7
coba.ntree25 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=25,
importance=TRUE, na.action=na.omit)
coba.ntree50 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=50,
importance=TRUE, na.action=na.omit)
coba.ntree250 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=250,
importance=TRUE, na.action=na.omit)

```


Lampiran 18. *Syntax Random Forest* (lanjutan)

```

coba.ntree500 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9 +
x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=500,
importance=TRUE, na.action=na.omit)
coba.ntree1000 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9
+ x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=1000,
importance=TRUE, na.action=na.omit)
coba.ntree1500 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9
+ x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=1500,
importance=TRUE, na.action=na.omit)
coba.ntree2000 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9
+ x10 + x11 + x12 , data=train, method = "rf", mtry=7, ntree=2000,
importance=TRUE, na.action=na.omit)
coba.ntree25
coba.ntree50
coba.ntree250
coba.ntree500
coba.ntree1000
coba.ntree1500
coba.ntree2000

#Random Forest dengan mtry optimum=7 dan ntree optimum=1000
print(coba.ntree1000)
plot(coba.ntree1000)
plot

#prediksi dan akurasi RF data train menggunakan mtry optimum dan ntree
optimum
pred.trainRF <- predict(coba.ntree1000)
pred.trainRF

#prediksi dan akurasi RF data testing menggunakan mtry optimum dan ntree
optimum
hasil.ntree1000 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 +x8 +x9
+ x10 + x11 + x12 , data=test, method = "rf", mtry=7, ntree=1000,
importance=TRUE, na.action=na.omit)
hasil.ntree1000
pred.testRF <- predict(hasil.ntree1000)
pred.testRF
# Importance variable RF
View(importance(hasil.ntree1000))
varImpPlot(hasil.ntree1000, pch=19, main = "Importance Variable")

```

Lampiran 19. Output Random Forest

```

> library(randomForest)
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
>
>
> #memanggil data
> mydata <- read.csv("E:/datakekeringan.csv")
> attach(mydata)
>
> #convert variabel x1, x4, x7, x8, x9, x10, x11, x12, dan y
kedalam bentuk variabel kategorik
> mydata$x1 <- factor(mydata$x1)
> mydata$x4 <- factor(mydata$x4)
> mydata$x7 <- factor(mydata$x7)
> mydata$x8 <- factor(mydata$x8)
> mydata$x9 <- factor(mydata$x9)
> mydata$x10 <- factor(mydata$x10)
> mydata$x11 <- factor(mydata$x11)
> mydata$x12 <- factor(mydata$x12)
> mydata$y <- factor(mydata$y)
>
> #training testing
> set.seed(123)
> train <- mydata[1:225, ]
> test <- mydata[226:300, ]
>
> #mencari mtry optimum dengan menggunakan ntree default=500
> coba.mtry2 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+ x8 + x9 + x10 + x11 + x12 , data=train, method = "rf",
mtry=2, ntree=500, importance=TRUE, na.action=na.omit)
> coba.mtry3 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6
+ x7 + x8 + x9 + x10 + x11 + x12 , data=train, method = "rf",
mtry=3, ntree=500, importance=TRUE, na.action=na.omit)
> coba.mtry7 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 + x6
+ x7 + x8 + x9 + x10 + x11 + x12 , data=train, method = "rf",
mtry=7, ntree=500, importance=TRUE, na.action=na.omit)
> coba.mtry2

Call:
randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+ x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 2, ntree = 500, importance = TRUE,
na.action = na.omit)
Type of random forest: classification
Number of trees: 500
No. of variables tried at each split: 2

OOB estimate of error rate: 32.89%
Confusion matrix:
  0  1 class.error
0 39  50  0.5617978
1 24 112  0.1764706

```

Lampiran 19. Output Random Forest (lanjutan)

```

> coba.mtry3

Call:
randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+ x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 3, ntree = 500, importance = TRUE,
na.action = na.omit)
Type of random forest: classification
Number of trees: 500
No. of variables tried at each split: 3

OOB estimate of error rate: 33.78%
Confusion matrix:
  0  1 class.error
0 42 47  0.5280899
1 29 107 0.2132353
> coba.mtry7

Call:
randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+ x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7, ntree = 500, importance = TRUE,
na.action = na.omit)
Type of random forest: classification
Number of trees: 500
No. of variables tried at each split: 7

OOB estimate of error rate: 32%
Confusion matrix:
  0  1 class.error
0 46 43  0.4831461
1 29 107 0.2132353

> #mencari ntree optimum dengan menggunakan mtry optimum=7
> coba.ntree25 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 +
x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=train, method =
"rf", mtry=7, ntree=25, importance=TRUE, na.action=na.omit)
> coba.ntree50 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 +
x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=train, method =
"rf", mtry=7, ntree=50, importance=TRUE, na.action=na.omit)
> coba.ntree250 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 +
x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=train, method =
"rf", mtry=7, ntree=250, importance=TRUE, na.action=na.omit)
> coba.ntree500 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5 +
x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=train, method =
"rf", mtry=7, ntree=500, importance=TRUE, na.action=na.omit)
> coba.ntree1000 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5
+ x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=train, method =
"rf", mtry=7, ntree=1000, importance=TRUE,
na.action=na.omit)
> coba.ntree25

```

Lampiran 19. Output Random Forest (lanjutan)

```

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+       x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7,       ntree = 25, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
    Number of trees: 25
No. of variables tried at each split: 7

      OOB estimate of error rate: 36.44%
Confusion matrix:
  0  1 class.error
0 46 43  0.4831461
1 39 97  0.2867647

> coba.ntree50

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+       x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7,       ntree = 50, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
    Number of trees: 50
No. of variables tried at each split: 7

      OOB estimate of error rate: 35.11%
Confusion matrix:
  0  1 class.error
0 39 50  0.5617978
1 29 107  0.2132353

> coba.ntree250

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+       x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7,       ntree = 250, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
    Number of trees: 250
No. of variables tried at each split: 7

      OOB estimate of error rate: 33.78%
Confusion matrix:
  0  1 class.error
0 45 44  0.4943820
1 32 104  0.2352941

```

Lampiran 19. Output Random Forest (lanjutan)

```

> coba.ntree500

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+             x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7, ntree = 500, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
  Number of trees: 500
No. of variables tried at each split: 7

  OOB estimate of error rate: 35.11%
Confusion matrix:
  0 1 class.error
0 43 46 0.5168539
1 33 103 0.2426471
> coba.ntree1000

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+             x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7, ntree = 1000, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
  Number of trees: 1000
No. of variables tried at each split: 7

  OOB estimate of error rate: 32.89%
Confusion matrix:
  0 1 class.error
0 47 42 0.4719101
1 32 104 0.2352941
> #Random Forest dengan mtry optimum=7 dan ntree
optimum=1000
> print(coba.ntree1000)

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+             x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7, ntree = 1000, importance = TRUE,
na.action = na.omit)
  Type of random forest: classification
  Number of trees: 1000
No. of variables tried at each split: 7

  OOB estimate of error rate: 32.89%
Confusion matrix:
  0 1 class.error
0 47 42 0.4719101
1 32 104 0.2352941

```

Lampiran 19. Output Random Forest (lanjutan)

```

> coba.ntree1500

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+           x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7,           ntree = 1500, importance = TRUE,
na.action = na.omit)
      Type of random forest: classification
      Number of trees: 1500
No. of variables tried at each split: 7

      OOB estimate of error rate: 35.11%
Confusion matrix:
      0  1 class.error
0 44 45  0.505618
1 34 102 0.250000
> coba.ntree2000

Call:
  randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+           x8 + x9 + x10 + x11 + x12, data = train, method =
"rf", mtry = 7,           ntree = 2000, importance = TRUE,
na.action = na.omit)
      Type of random forest: classification
      Number of trees: 2000
No. of variables tried at each split: 7

      OOB estimate of error rate: 35.56%
Confusion matrix:
      0  1 class.error
0 42 47  0.5280899
1 33 103 0.2426471
> plot(coba.ntree1000)
> plot
function (x, y, ...)
UseMethod("plot")
<bytecode: 0x0000000009410790>
<environment: namespace:graphics>
> #prediksi dan akurasi RF data testing menggunakan mtry
optimum dan ntree optimum
> hasil.ntree1000 <- randomForest(y ~ x1 + x2 + x3 + x4 + x5
+ x6 + x7 +x8 +x9 + x10 + x11 + x12 , data=test, method =
"rf", mtry=7, ntree=1000, importance=TRUE,
na.action=na.omit)

```

Lampiran 19. Output Random Forest (lanjutan)

```
> hasil.ntree1000
```

```
Call:
```

```
randomForest(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7
+ x8 + x9 + x10 + x11 + x12, data = test, method =
"rf", mtry = 7, ntree = 1000, importance = TRUE,
na.action = na.omit)
```

```
      Type of random forest: classification
```

```
      Number of trees: 1000
```

```
No. of variables tried at each split: 7
```

```
      OOB estimate of error rate: 24%
```

```
Confusion matrix:
```

```
      0 1 class.error
0 10 10  0.5000000
1  8 47  0.1454545
```

```
> # Importance variable RF
```

```
> View(importance(hasil.ntree1000))
```

	0	1	MeanDecreaseAccuracy	MeanDecreaseGini
x1	2.0008788	-2.2386800	-0.2990974	0.07142362
x2	5.3977052	3.7594006	6.2111833	4.34575133
x3	5.6575456	10.7630289	11.6950941	4.09684900
x4	29.6947652	25.3271595	35.6873713	7.34756954
x5	0.8243823	6.5379464	6.2265036	2.54369019
x6	-0.9187041	-1.0413016	-1.2160142	3.93755917
x7	-1.6932787	-1.7264303	-2.2087793	0.34386499
x8	1.4292466	-0.4994419	0.9661891	0.45423284
x9	1.0078899	-4.5134292	-3.0154562	0.51784341
x10	-0.7063702	-3.0910636	-2.5017846	0.68831655
x11	-4.4552919	13.1597212	9.5803239	3.22187284
x12	3.7784160	2.0865901	3.9494320	1.32827985

Halaman ini sengaja dikosongkan

BIODATA PENULIS



Penulis bernama lengkap Pratiwi Penta Atrivi. Lahir di Surabaya tanggal 22 Desember 1995. Penulis merupakan anak terakhir dari empat bersaudara. Penulis bertempat tinggal di Jalan Bulak Bnateng Baru Gang Cempaka No.29 Surabaya. Penulis menyelesaikan pendidikan di TK Al Sari Surabaya tahun 2002, SDN Bulak Banteng 1/263 Surabaya tahun 2008, SMPN 15 Surabaya tahun 2011, SMAN 9 Surabaya tahun 2014. Setelah lulus SMA, penulis mengikuti seleksi penerimaan mahasiswa baru dan diterima di program Diploma III melalui jalur seleksi reguler pada tahun 2014. Selama masa perkuliahan penulis memiliki pengalaman kerja sebagai surveyor di PT Mitra Pinasthika Mulia (MPM). Selain itu, penulis aktif mengikuti organisasi Himpunan Mahasiswa saat menempuh Diploma Statistika ITS sebagai staff dan Sekretaris Departemen PSDM (Pengembangan Sumber Daya Mahasiswa). Selain berpengalaman dalam organisasi, penulis juga aktif dalam kepanitiaan diantaranya MAKRAB (Malam Keakraban) FMIPA ITS, OKKBK (Orientasi Keprofesian dan Kompetensi Berbasis Kurikulum), STATION (*Statistics Competition*), CERITA (Cerdas Bersama Statistika), LKMM TD, Pelepasan Wisuda Departemen Statistika Bisnis, ketua panitia PENTAKASI (Pengenalan HIMADATA-ITS, IKASTA-ITS dan IHMSI) dan lain-lain. Setelah menyelesaikan jenjang Diploma, penulis kemudian melanjutkan ke jenjang S1 ITS melalui Lintas Jalur pada Tahun 2017. Saat menempuh perkuiahkan jenjang S1 Statistika ITS, penulis pernah melakukan kerja praktek di Badan Penanggulangan Bencana Daerah Provinsi Jawa Timur. Penulis dapat dihubungi melalui email pratiwi.pentaatrivi@gmail.com apabila pembaca memiliki kritik, saran maupun berdiskusi lebih lanjut mengenai tugas akhir ini.