

**PREDIKSI KEJADIAN HUJAN
MENGUNAKAN METODE *ENSEMBLE*
DENGAN ALGORITMA *FUZZY NAÏVE BAYES*, *NAÏVE BAYES*
DAN *DECISION TREE*
BERDASARKAN *RESAMPLING BOOTSTRAP***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana di
Jurusan Matematika pada Fakultas MIPA**

Oleh :

AYU DWI PANGESTI

08011181924004



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2023

LEMBAR PENGESAHAN

**PREDIKSI KEJADIAN HUJAN
MENGUNAKAN METODE *ENSEMBLE*
DENGAN ALGORITMA *FUZZY NAÏVE BAYES*, *NAÏVE BAYES*
DAN *DECISION TREE* BERDASARKAN *RESAMPLING BOOTSTRAP***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana di
Jurusan Matematika pada Fakultas MIPA**

Oleh

**AYU DWI PANGESTI
08011181924004**

Pembimbing Kedua



**Drs. Endro Setyo Cahyono, M.Si
NIP. 196409261990021002**

**Indralaya, Mei 2023
Pembimbing Utama**



**Dr. Yulia Resti, S.Si., M.Si
NIP. 197307191997022001**

**Mengetahui,
Ketua Jurusan Matematika**



**Drs. Sugandi Yahdin, M.M
NIP. 19580727198603 1003**

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Ayu Dwi Pangesti
NIM : 08011181924004
Fakultas/Jurusan : MIPA/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan sarjana satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat dalam skripsi ini berasal dari penulis lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, Mei 2023



Penulis

HALAMAN PERSEMBAHAN

MOTTO

“Barangsiapa bertakwa kepada Allah niscaya Dia akan membukakan jalan keluar baginya, dan Dia memberinya rezeki dari arah yang tidak disangka-sangkanya.

Dan barangsiapa bertawakal kepada Allah, niscaya Allah akan mencukupkan (keperluan)nya. Sesungguhnya Allah melaksanakan urusan-Nya. Sungguh, Allah telah mengadakan ketentuan bagi setiap sesuatu.”

Q.S At-Talaq (65) : 3

“Cukuplah Allah (menjadi penolong) bagi kami dan Dia sebaik baik pelindung.”

Skripsi ini kupersembahkan kepada:

- 1. Allah SWT**
- 2. Orangtuaku**
- 3. Saudaraku**
- 4. Keluarga Besar**
- 5. Dosen dan Guruku**
- 6. Almamater**
- 7. Sahabat dan Temanku**

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Dengan mengucapkan Alhamdulillah, segala puji bagi Allah Subhanahu Wa Ta'ala atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini yang berjudul **“Prediksi Kejadian Hujan Menggunakan Metode *Ensemble* dengan Algoritma *Fuzzy Naïve Bayes*, *Naïve Bayes* dan *Decision Tree* Berdasarkan *Resampling Bootstrap*”** dapat berjalan dengan baik. Shalawat beserta salam senantiasa tercurahkan kepada baginda kita Nabi Muhammad SAW beserta keluarga, sahabat dan para pengikutnya hingga akhir zaman. Skripsi ini merupakan salah satu persyaratan yang harus dipenuhi penulis untuk meraih gelar Sarjana Sains Bidang Studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Dengan segala hormat dan kerendahan hati penulis mengucapkan terima kasih yang sebesar-besarnya dan penghargaan yang tulus penulis sampaikan kepada kedua orang tua tercinta, yaitu **Bapak Harnadi** dan **Ibu Susanti** yang telah menuntut, mendidik, mengajari, menasehati, memberi semangat dan tidak pernah lelah berdo'a yang terbaik untuk anaknya. Penulis juga menyampaikan ucapan terima kasih dan penghargaan kepada:

1. Bapak **Drs. Sugandi Yahdin, M.M** selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya sekaligus Dosen Pembahas I yang telah memberikan waktu serta saran dan tanggapannya yang bermanfaat bagi penulis dalam mengerjakan skripsi.

2. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya yang telah membimbing dan mengarahkan urusan akademik kepada penulis.
3. Ibu **Dr. Yulia Resti, S.Si., M.Si** dan Bapak **Drs. Endro Setyo Cahyono, M.Si** selaku Dosen Pembimbing Utama dan Dosen Pembimbing Kedua yang telah bersedia meluangkan waktu, tenaga dan pikiran untuk memberikan bimbingan dan pengarahan dengan penuh perhatian, pengertian dan kesabaran sehingga skripsi ini dapat diselesaikan.
4. Bapak **Dr. Bambang Suprihatin, S.Si., M.Si** selaku Dosen Pembahas II yang telah memberikan tanggapan, kritik dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
5. Ibu **Dra. Ning Eliyati, M.Pd** selaku Ketua Seminar Skripsi yang telah memberikan arahan, saran serta mengatur jalannya seminar sehingga dapat berjalan dengan baik.
6. Ibu **Dr. Yuli Andriani, S.Si., M.Si** selaku Sekretaris Seminar Skripsi yang telah membantu dan memberikan catatan serta masukan yang sangat berguna bagi penulis.
7. Bapak **Drs. Robinson Sitepu, M.Si** selaku Dosen Pembimbing Akademik yang telah memberikan semangat, motivasi dan pengarahan kepada penulis selama masa perkuliahan.
8. **Seluruh Dosen di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam** yang telah memberikan ilmu yang bermanfaat, nasihat serta motivasi selama penulis menjalani perkuliahan.

9. Bapak **Irwansyah** dan Ibu **Hamidah** selaku Pegawai Tata Usaha Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah membantu penulis selama perkuliahan.
10. Keluarga besarku terutama kakakku **Wahyu Herliyanto** terima kasih atas do'a, kasih sayang, semangat, motivasi serta dukungan yang telah diberikan kepada penulis.
11. Sahabat-sahabatku **Khoirotun Nisa, Leliani, Nadia Vionica, Niluh Mutiara Komala Ayu, Silvi Lisandi** dan adikku **Vira Ardhana Andriyani** terima kasih untuk saling menguatkan, mengajarkan, memberikan kesan yang tak terlupakan dan memberikan bantuannya selama proses penyelesaian skripsi ini, sukses dan sehat selalu untuk kalian.
12. Teman seperjuangan skripsi **Chika, Sinta, Elisa, Sherly, Vina, Uci, Tia, Alga** serta seluruh teman-teman angkatan 2019 terima kasih untuk saling menguatkan, mengajarkan dan kebersamaannya, sehat dan sukses selalu untuk kalian semua.
13. Kak **Nurafni Rahayu Khotimah** terima kasih atas arahan dan bantuan dalam proses pengerjaan skripsi ini.
14. Kakak-kakak tingkat angkatan 2017, 2018 dan adik-adik tingkat angkatan 2020 dan 2021 atas bantuan selama perkuliahan.
15. Semua **pihak** yang tidak dapat penulis sebutkan satu persatu yang telah memberikan bantuan dalam menyelesaikan skripsi ini hanya ucapan terima kasih yang dapat penulis berikan. Semoga segala kebaikan yang diberikan mendapatkan balasan dari Allah Subhanahu Wa Ta'ala.

16. *Last but not least, i would like to thank myself for successfully going through the long process of studying until the completion of this thesis. Thank you for being patient, persisting and always being healthy in every process of this thesis or this process of self-development.*

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam dan semua pihak yang memerlukan.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, Mei 2023

Penulis

USING RAIN EVENTS PREDICTION ENSEMBLE METHOD WITH FUZZY NAÏVE BAYES, NAÏVE BAYES AND DECISION TREE ALGORITHMS BASED BOOTSTRAP RESAMPLING

By:

AYU DWI PANGESTI

08011181924004

ABSTRACT

Precise weather predictions are needed in various fields including agriculture, tourism, aviation, shipping and plantations. The most predictable weather element is rain. This study aims to predict rain events using the ensemble method with fuzzy naïve Bayes, naïve Bayes and decision tree algorithms based on bootstrap resampling. This study uses weather datasets in Australia sourced from the Kaggle website, totaling 145460 data. The level of accuracy obtained in predicting rain events using the fuzzy naïve Bayes method produces an average value of accuracy, precision, recall and fscore of 78.19%, 50%, 51.34% and 50.66%. The naïve Bayes method produces an average value of accuracy, precision, recall and fscore of 77.13%, 48.63%, 37.77% and 42.52%. The decision tree method produces an average value of accuracy, precision, recall and fscore of 76.66%, 48%, 50.90% and 49.41%. The ensemble method produces an average value of accuracy, precision, recall and fscore of 79.29%, 54.48%, 45.61% and 49.65%. The results show that the ensemble method produces better prediction accuracy because it produces higher accuracy and precision values compared to the other three methods. However, the naïve Bayes fuzzy method has a better recall and fscore than the other three methods.

Keyword: Rain Prediction, Ensemble, Fuzzy Naïve Bayes, Naïve Bayes, Decision Tree, Resampling Bootstrap

**PREDIKSI KEJADIAN HUJAN
MENGUNAKAN METODE *ENSEMBLE*
DENGAN ALGORITMA *FUZZY NAÏVE BAYES*, *NAÏVE BAYES*
DAN *DECISION TREE*
BERDASARKAN *RESAMPLING BOOTSTRAP***

Oleh:

AYU DWI PANGESTI

08011181924004

ABSTRAK

Prediksi cuaca yang tepat sangat dibutuhkan dalam berbagai bidang diantaranya pertanian, pariwisata, penerbangan, pelayaran dan perkebunan. Unsur cuaca yang paling sering di prediksi adalah hujan. Penelitian ini bertujuan untuk memprediksi kejadian hujan menggunakan metode *Ensemble* dengan algoritma *Fuzzy Naïve Bayes*, *Naïve Bayes* dan *Decision Tree* berdasarkan *Resampling Bootstrap*. Penelitian ini menggunakan *dataset* cuaca di Australia yang bersumber dari *website Kaggle*, berjumlah 145.460 data. Tingkat ketepatan yang diperoleh dalam memprediksi kejadian hujan menggunakan metode *Fuzzy Naïve Bayes* menghasilkan nilai rata-rata akurasi, presisi, *recall* dan *fscore* sebesar 78.19%, 50%, 51.34% dan 50.66%. Metode *Naïve Bayes* menghasilkan nilai rata-rata akurasi, presisi, *recall* dan *fscore* sebesar 77.13%, 48.63%, 37.77% dan 42.52%. Metode *Decision Tree* menghasilkan nilai rata-rata akurasi, presisi, *recall* dan *fscore* sebesar 76.66%, 48%, 50.90% dan 49.41%. Metode *Ensemble* menghasilkan nilai rata-rata akurasi, presisi, *recall* dan *fscore* sebesar 79.29%, 54.48%, 45.61% dan 49.65%. Hasil penelitian menunjukkan bahwa metode *Ensemble* menghasilkan ketepatan prediksi lebih baik karena menghasilkan nilai akurasi dan presisi yang lebih tinggi dibandingkan dengan ketiga metode lainnya. Namun pada metode *Fuzzy Naïve Bayes* memiliki nilai *recall* dan *fscore* yang lebih baik dibandingkan dengan ketiga metode lainnya.

Kata kunci: Prediksi Hujan, *Ensemble*, *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree*, *Resampling Bootstrap*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSEMBAHAN	iii
KATA PENGANTAR.....	iv
ABSTRACT.....	viii
ABSTRAK	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	7
BAB II TINJAUAN PUSTAKA.....	8
2.1 Hujan	8
2.2 <i>Statistical Machine Learning</i>	8
2.3 Prediksi.....	9
2.4 Diskritisasi Data.....	10
2.5 <i>Resampling Bootstrap</i>	10
2.6 Peluang.....	12
2.7 Metode <i>Naïve Bayes</i>	12
2.8 <i>Laplace Smoothing</i>	15
2.9 <i>Decision Tree</i>	15
2.10 Himpunan <i>Fuzzy</i>	17
2.11 Fungsi Keanggotaan <i>Fuzzy</i>	17
2.12 Metode <i>Fuzzy Naïve Bayes</i>	20
2.13 Metode <i>Ensemble</i>	21
2.14 <i>Confusion Matrix</i>	22
BAB III METODOLOGI PENELITIAN	24
3.1 Tempat Penelitian.....	24
3.2 Waktu Penelitian	24
3.3 Metode Penelitian.....	24
BAB IV HASIL DAN PEMBAHASAN	27
4.1 Deskripsi Data.....	27
4.2 Prediksi Menggunakan Metode <i>Fuzzy Naïve Bayes</i>	27

4.2.1 Menentukan Himpunan Universal	28
4.2.2 Menentukan Nilai Keanggotaan <i>Fuzzy</i>	38
4.2.3 Partisi Data	42
4.2.4 <i>Resampling Bootstrap</i>	43
4.2.5 Menghitung Nilai Probabilitas <i>Prior</i>	43
4.2.6 Menghitung Nilai Probabilitas <i>Likelihood</i> Y dengan Syarat X.....	44
4.2.7 Menghitung Nilai <i>Posterior</i>	46
4.2.8 Menghitung Nilai Ketepatan Prediksi Yaitu Akurasi, Presisi, <i>Recall</i> Dan <i>Fscore</i>	47
4.3 Prediksi Menggunakan Metode <i>Naïve Bayes</i>	49
4.3.1 Diskritisasi Data	49
4.3.2 Partisi Data	52
4.3.3 <i>Resampling Bootstrap</i>	53
4.3.4 Menghitung Nilai Probabilitas <i>Prior</i>	53
4.3.5 Menghitung Nilai Probabilitas <i>Likelihood</i> Y dengan Syarat X.....	54
4.3.6 Menghitung Nilai <i>Posterior</i>	56
4.3.7 Menghitung Nilai Ketepatan Prediksi Yaitu Akurasi, Presisi, <i>Recall</i> Dan <i>Fscore</i>	57
4.4 Prediksi Menggunakan Metode <i>Decision Tree</i>	59
4.4.1 Menghitung Nilai Ketepatan Prediksi Yaitu Akurasi, Presisi, <i>Recall</i> Dan <i>Fscore</i>	63
4.5 Metode <i>Ensemble</i>	65
4.6 Analisis Hasil	67
BAB V KESIMPULAN DAN SARAN	69
5.1 Kesimpulan	69
5.2 Saran	69
DAFTAR PUSTAKA	71
LAMPIRAN.....	75

DAFTAR TABEL

Tabel 2.1 Metode <i>Ensemble Majority Voting</i>	21
Tabel 2.2 <i>Confusion Matrix</i>	22
Tabel 4.1 Dataset Cuaca.....	27
Tabel 4.2 Interval himpunan <i>fuzzy</i> variabel prediktor X_2	28
Tabel 4.3 Himpunan <i>fuzzy</i> variabel prediktor X_2	29
Tabel 4.4 Interval himpunan <i>fuzzy</i> variabel prediktor X_3	31
Tabel 4.5 Himpunan <i>fuzzy</i> variabel prediktor X_3	31
Tabel 4.6 Interval himpunan <i>fuzzy</i> variabel prediktor X_4	33
Tabel 4.7 Himpunan <i>fuzzy</i> variabel prediktor X_4	33
Tabel 4.8 Interval himpunan <i>fuzzy</i> variabel prediktor.....	35
Tabel 4.9 Himpunan <i>fuzzy</i> variabel prediktor	37
Tabel 4.10 Nilai keanggotaan terbesar variabel prediktor X_2	39
Tabel 4.11 Nilai keanggotaan terbesar variabel prediktor X_3	40
Tabel 4.12 Nilai keanggotaan terbesar variabel prediktor X_4	41
Tabel 4.13 Nilai keanggotaan terbesar variabel prediktor	41
Tabel 4.14 Data Latih.....	42
Tabel 4.15 Data Uji	42
Tabel 4.16 Data <i>Latih Resampling Bootstrap</i>	43
Tabel 4.17 Nilai probabilitas <i>likelihood</i>	45
Tabel 4.18 Hasil prediksi metode <i>Fuzzy Naïve Bayes</i>	47
Tabel 4.19 <i>Confusion matrix</i> metode <i>Fuzzy Naïve Bayes</i>	47
Tabel 4.20 <i>Range</i> variabel prediktor	50
Tabel 4.21 Hasil diskritisasi data	51
Tabel 4.22 Data Latih.....	52
Tabel 4.23 Data Uji	52
Tabel 4.24 Data Latih <i>Resampling Bootstrap</i>	53
Tabel 4.25 Nilai probabilitas <i>likelihood</i>	55
Tabel 4.26 Hasil prediksi metode <i>Naïve Bayes</i>	57
Tabel 4.27 <i>Confusion matrix</i> metode <i>Naïve Bayes</i>	57
Tabel 4.28 Nilai <i>entropy</i> variabel prediktor	60
Tabel 4.29 Nilai <i>gain</i> variabel prediktor	61
Tabel 4.30 Hasil prediksi menggunakan metode <i>decision tree</i>	63
Tabel 4.31 <i>Confusion matrix</i> metode <i>decision tree</i>	63
Tabel 4.32 Hasil prediksi menggunakan metode <i>ensemble</i>	65
Tabel 4.33 <i>Confusion matrix</i> metode <i>ensemble</i>	66
Tabel 4.34 Perbandingan ketepatan hasil prediksi	67

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Tahapan <i>Resampling Bootstrap</i>	11
Gambar 2.2 Representasi Kurva Linier Naik.....	18
Gambar 2.3 Representasi Kurva Linier Turun.....	19
Gambar 2.4 Representasi Kurva Segitiga	19
Gambar 2.5 Representasi Kurva Trapesium	20
Gambar 4.1 Pohon Keputusan <i>Decision Tree</i> ID3.....	62

DAFTAR LAMPIRAN

Lampiran 1 Batasan untuk masing-masing variabel prediktor	76
---	----

BAB I

PENDAHULUAN

1.1 Latar Belakang

Prediksi cuaca sangat membantu banyak aktivitas masyarakat. Dengan prediksi cuaca memungkinkan seseorang untuk merencanakan dan mengambil tindakan pencegahan terhadap berbagai bencana alam, seperti banjir dan angin topan sehingga dapat meminimalisir dampak yang terjadi. Prediksi cuaca yang tepat sangat dibutuhkan dalam berbagai bidang seperti pertanian, pariwisata, penerbangan, pelayaran, perkebunan, kehutanan, pembangunan gedung, penataan wilayah dan kesehatan (T.P & Sari, 2017).

Prediksi cuaca merupakan suatu studi yang dapat dilakukan untuk mendapatkan perkiraan keadaan cuaca di masa mendatang berdasarkan pada data historis dari keadaan cuaca yang telah terjadi sebelumnya (Adnyana et al., 2019). Kondisi cuaca di suatu tempat dipengaruhi oleh beberapa faktor seperti suhu udara, kelembaban udara, intensitas cahaya matahari, arah angin, kecepatan angin, dan sebagainya. Sehingga dengan melihat faktor-faktor tersebut dapat diprediksi cuaca yang akan terjadi pada keesokan harinya (T.P & Sari, 2017). Unsur cuaca yang paling sering di prediksi adalah hujan.

Dengan pesatnya teknologi yang berkembang saat ini, prediksi hujan dapat dilakukan dengan menggunakan *statistical machine learning*. *Statistical machine learning* adalah suatu metode yang digunakan untuk mendapatkan hasil prediksi dari klasifikasi data yang tepat. Statistika memiliki fungsi utama yaitu mempelajari teori-teori statistik yang dapat digunakan untuk menginterpretasikan dan

menjelaskan model, sedangkan *machine learning* memanfaatkan data untuk membuat model statistik. Model tersebut pada umumnya digunakan sistem untuk melakukan prediksi di masa depan berdasarkan data masa lalu yang diinputkan atau mempelajari pola yang terdapat dalam data. *Statistical machine learning* mencakup banyak metode seperti *Logistic Regression*, *Classification and Regression Trees*, *Support Vector Machine* dan *K-Nearest Neighbor* (James *et al.*, 2013).

Salah satu metode yang dapat digunakan dalam memprediksi kejadian hujan yakni metode *Fuzzy Naïve Bayes*. Metode *Fuzzy Naïve Bayes* merupakan metode klasifikasi gabungan yang terdiri dari teori himpunan *fuzzy* dan klasifikasi *Naïve Bayes* (Insani, 2017). Teori himpunan fuzzy adalah kerangka matematika untuk merepresentasikan ketidakpastian, ambiguitas, ketidaklengkapan, ketidaktepatan informasi dan kebenaran parsial. Penentuan nilai kriteria subjektif dengan proses *fuzzy* memberikan hasil yang sangat baik (Ichwan *et al.*, 2018). Sedangkan metode *naïve bayes* adalah metode klasifikasi statistik yang dapat memprediksi probabilitas suatu kelas yang mendukung teorema Bayes (Putri *et al.*, 2021).

Penelitian terdahulu oleh Insani (2017) dengan judul Sistem Pakar Deteksi Dini Penyakit Stroke Menggunakan Metode *Fuzzy Naïve Bayes*, hasil pengujian akurasi terhadap 30 data kasus uji menghasilkan persentase sebesar 85,6%. Penelitian lain dilakukan oleh Putri *et al* (2021) dengan judul *Sentiment Analysis Using Fuzzy Naïve Bayes Classifier On Covid-19* menggunakan data *training* dan *testing* 70% : 30% menghasilkan akurasi sebesar 83,1% berdasarkan 1199 data *tweet* yang terdiri dari 360 data uji dan 839 data latih. Penelitian terdahulu

menggunakan metode *Naïve Bayes* salah satunya oleh Gunawan dan Fernando (2021) yang berjudul Sistem Pakar Diagnosa Penyakit Kulit Pada Kucing Menggunakan Metode *Naive Bayes* Berbasis *Web*. Dari Pengujian 15 data rekam medis dokter menghasilkan tingkat akurasi dalam mendiagnosa penyakit kulit pada kucing sebesar 80%.

Algoritma lain yang dapat digunakan untuk memprediksi kejadian hujan adalah *decision tree*. *Decision tree* adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Salah satu algoritma yang termasuk ke dalam *decision tree* adalah algoritma *Iterative Dichotomiser Tree* (ID3) (Choina *et al.*, 2020). Penelitian terdahulu menggunakan algoritma ID3 salah satunya oleh T.P & Sari (2017) dengan judul Penerapan *Data Mining* Untuk Prakiraan Cuaca Di Kota Malang Menggunakan Algoritma *Iterative Dichotomiser Tree* (ID3), menghasilkan akurasi sebesar 73.74%. Penelitian lain dilakukan oleh Karo *et al* (2022) dengan judul Klasifikasi Kebakaran Hutan Menggunakan *Feature Selection* Dengan Algoritma K-NN, *Naive Bayes* Dan ID3, menghasilkan performa algoritma ID3 yang lebih unggul dibandingkan dengan algoritma lainnya dengan akurasi 65.83%, *precision* 67.4%, *recall* 67.02 % dan F1 67.21%.

Namun pada penerapan metode klasifikasi tunggal, terkadang masih memiliki kelemahan dalam melakukan keakuratan prediksi. Maka membutuhkan suatu cara untuk meningkatkan akurasi dari hasil pembelajaran klasifikasi yaitu metode *ensemble* (Aziz & Fanani, 2023). Metode *ensemble* melatih sejumlah *classifier* atau *base learner* dan kemudian mengkombinasikan hasilnya (Khadijah & Kusumaningrum, 2019). *Base learner* yang digunakan dapat berupa algoritma *fuzzy*

naïve bayes, *naïve bayes*, dan *decision tree*. Penelitian terdahulu oleh Indriati & Kusyanti (2019) dengan judul Metode *Ensemble Classifier* Untuk Mendeteksi Jenis *Attention Deficit Hyperactivity Disorder* (ADHD) pada Anak Usia Dini, mengidentifikasi ADHD menggunakan metode *Ensemble Classifier* yaitu dengan menggunakan hasil penentuan *class* dari *single classifier* kemudian dengan menggunakan *majority voting* menentukan *class* berdasarkan *class* mayoritas, menghasilkan akurasi tertinggi dari *ensemble classifier* yaitu 95% dengan nilai *K* optimal yaitu $K = 10$.

Agar metode klasifikasi yang digunakan dapat berjalan lebih baik, dapat digunakan *resampling bootstrap*. Metode *resampling bootstrap* merupakan metode yang digunakan untuk mengestimasi suatu distribusi populasi yang tidak diketahui dengan menggunakan distribusi empiris yang diperoleh dari proses pengambilan sampel ulang dari sampel asli dengan ukuran sama dengan ukuran sampel asli dan dilakukan dengan pengembalian (Atinri *et al.*, 2016). Penelitian terdahulu dengan judul Metode *Sample Bootstrapping* Untuk Meningkatkan Performa Algoritma *Naive Bayes* Pada Citra Tunggal *Pap Smear* oleh Dewi & Sariasih (2019), menunjukkan bahwa metode *sample bootstrapping* dapat meningkatkan nilai akurasi tujuh kelas menjadi 85,24% dan 93,24% untuk nilai akurasi dua kelas. Penelitian lain dilakukan oleh Khotimah *et al* (2022) dengan judul Implementasi *Bootstrap Sampling* Pada Metode *Ensemble* Berdasarkan Metode *Naïve Bayes*, *Random Forest* dan Regresi Logistik Biner Untuk Prediksi Kejadian Hujan, menunjukkan bahwa metode *ensemble* menghasilkan rata-rata akurasi dan *fscore*

lebih tinggi dibandingkan dengan ketiga metode lain yaitu sebesar 82.83% dan 53.98%.

Penelitian terdahulu mengenai prediksi kejadian hujan menggunakan *dataset* cuaca di Australia yang bersumber dari *website Kaggle* oleh Saputra *et al* (2021) dengan judul *Application Of Data Mining For Rainfall Prediction Classification In Australia With Decision Tree Algorithm And C5.0 Algorithm*, didapatkan akurasi paling tinggi sebesar 87.35% untuk algoritma *decision tree* dan akurasi sebesar 86.85% untuk algoritma C5.0.

Berdasarkan penelitian-penelitian terdahulu, peneliti mencoba untuk memprediksi kejadian hujan pada *dataset* cuaca di Australia yang bersumber dari *website Kaggle* menggunakan metode *Ensemble* dengan algoritma *Fuzzy Naïve Bayes*, *Naïve Bayes* dan *Decision Tree* berdasarkan *Resampling Bootstrap*, serta tingkat ketepatan prediksi pada penelitian ini dibatasi oleh nilai akurasi, presisi, *recall* dan *f-score*.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah:

1. Menentukan tingkat ketepatan dalam memprediksi kejadian hujan menggunakan metode *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree* dan *Ensemble* berdasarkan *Resampling Bootstrap*.
2. Membandingkan tingkat ketepatan dalam memprediksi kejadian hujan pada metode *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree* dan *Ensemble* berdasarkan *Resampling Bootstrap*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Penelitian ini menggunakan *dataset* cuaca di Australia yang bersumber dari *website Kaggle*. Data ini berjumlah 145.460 data, terdiri dari 21 variabel prediktor yaitu *Location*, *MinTemp*, *MaxTemp*, *Rainfall*, *Evaporation*, *Sunshine*, *WindGustDir*, *WindGustSpeed*, *WindDir9am*, *WindDir3pm*, *WindSpeed9am*, *WindSpeed3pm*, *Humidity9am*, *Humidity3pm*, *Pressure9am*, *Pressure3pm*, *Cloud9am*, *Cloud3pm*, *Temp9am*, *Temp3pm* dan *RainToday* serta 1 variabel respon yaitu *RainTomorrow*.
2. Menggunakan 4 fungsi keanggotaan *fuzzy* yaitu kurva linier naik, kurva linier turun, kurva segitiga dan kurva trapesium.
3. Atribut dari variabel prediktor yang digunakan terdiri dari 3, 5 dan 7 kategori. Pada 3 kategori terdiri dari rendah, sedang dan tinggi. Pada 5 kategori terdiri dari sangat rendah, rendah, sedang, tinggi dan sangat tinggi. Sedangkan pada 7 kategori terdiri dari sangat rendah, agak rendah, rendah, sedang, agak tinggi, tinggi dan sangat tinggi.
4. Tingkat ketepatan prediksi pada penelitian ini dibatasi oleh nilai akurasi, presisi, *recall* dan *fscore*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Memperoleh tingkat ketepatan dalam memprediksi kejadian hujan menggunakan metode *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree* dan *Ensemble* berdasarkan *Resampling Bootstrap*.
2. Memperoleh perbandingan tingkat ketepatan dalam memprediksi kejadian hujan pada metode *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree* dan *Ensemble* berdasarkan *Resampling Bootstrap*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini sebagai berikut:

1. Sebagai sarana pembelajaran dan memperluas pengetahuan bagi penulis maupun pembaca mengenai prediksi hujan menggunakan metode *Ensemble* dengan algoritma *Fuzzy Naïve Bayes*, *Naïve Bayes* dan *Decision Tree* berdasarkan *Resampling Bootstrap*.
2. Sebagai bahan referensi bagi peneliti lain terkait prediksi hujan, *Fuzzy Naïve Bayes*, *Naïve Bayes*, *Decision Tree*, *Ensemble* dan *Resampling Bootstrap*.

DAFTAR PUSTAKA

- Adnyana, I. N. T., Wijaya, I. G. P. S., & Albar, M. A. (2019). Jaringan Syaraf Tiruan Backpropagation Untuk Peramalan Suhu Minimum dan Maksimum. *J-Cosine*, 3(2), 127–136.
- Aldrian, E., Karmini, M., & Budiman. (2011). Adaptasi dan Mitigasi Perubahan Iklim di Indonesia. Jakarta: Pusat Perubahan Iklim dan Kualitas Udara Kedeputan Bidang Klimatologi (BMKG).
- Anas., Tempola, F., & Khairan, A. (2019). Hybrid fuzzy dan Naive Bayes Dalam Penentuan Status UKT. *Jurnal PROtek Vol*, 6(1).
- Atinri, O., Yoza, H., & Asdi, Y. (2016). Penentuan Ukuran Contoh Dan Replikasi Bootstrap Untuk Menduga Model Regresi Linier Sederhana. *Jurnal Matematika*, 3(2), 53–61.
- Aziz, M. I., & Fanani, A. Z. (2023). Analisis Metode Ensemble Pada Klasifikasi Penyakit Jantung Berbasis Decision Tree. 7, 1–12.
- Choina, I., Aulia, R., & Zakir, A. (2020). Penerapan Algoritma ID3 Untuk Menyeleksi Pegawai Kontrak Di Kantor Pengadilan Kota Langsa. *CESS (Journal of Computer Engineering, System and Science)*, 5(1), 47.
- Darwanto., & Dinata, K.B. (2021). *Pengantar Teori Peluang*. Lampung: UMKO Publishing.
- Davvaz, B., Mukhlash, I., & Soleha, S. (2021). Himpunan Fuzzy dan Rough Sets. *Limits: Journal of Mathematics and Its Applications*, 18(1), 79.
- Dewi, Y. N., & Sariasih, F. A. (2019). Metode Sample Bootstrapping Untuk Meningkatkan Performa Algoritma Naive Bayes Pada Citra Tunggal Pap Smear. *Jurnal Teknik Informatika*, 12(1), 1-10.
- Efendi, M. S., & Wibawa, H. A. (2018). Prediksi penyakit diabetes menggunakan algoritma ID3 dengan pemilihan atribut terbaik. *JUITA: Jurnal Informatika*, 6(1), 29-35.
- Endah, F., & Encis, I. S. Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim Amni). *Prosiding Kemaritiman 2021*.
- Gunawan, I., & Fernando, Y. (2021). Sistem Pakar Diagnosa Penyakit Kulit Pada Kucing Menggunakan Metode Naive Bayes Berbasis Web. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(2), 239-247.

- Gunawan, M. I., Sugiarto, D., & Mardianto, I. (2020). Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Search pada Algoritma Logistic Regression. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(3), 280.
- Ichwan, M., Husada, M. G., & FH, G. N. (2016). Penerapan Fuzzy Logic Tsukamoto pada Pembangunan Kandang Ayam Pintar. *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, 1(2), 11-14.
- Indriati., & Kusyanti, A. (2019). Metode Ensemble Classifier Untuk Mendeteksi Jenis Attention Deficit Hyperactivity Disorder (Adhd) Pada Anak Usia Dini. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 6(3).
- Insani, S. (2017). *Sistem Pakar Deteksi Dini Penyakit Stroke Menggunakan Metode Fuzzy Naive Bayes* (Doctoral dissertation, Universitas Brawijaya).
- Irawan, L., Hasibuan, L. H., & Fauzi, F. (2020). Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma ID3 Menggunakan Aplikasi Data Mining Orange. *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, 14(2), 189-201.
- James, Gareth dkk. 2013. *An Introduction to Statistical Learning with Applications in R*. New York: Springer.
- Karo, I. M. K., Amalia, S. N., & Septiana, D. (2022). Klasifikasi Kebakaran Hutan Menggunakan Feature Selection dengan Algoritma K-NN, Naive Bayes dan ID3. *Journal of Software Engineering, Information and Communication Technology (SEICT)*, 3(1), 121-126.
- Khadijah, K., & Kusumaningrum, R. (2019). Ensemble Classifier untuk Klasifikasi Kanker Payudara. *It Journal Research and Development*, 4(1), 61–71.
- Khotimah, N. R., Resti, Y., & Kresnawati, E. S. (2022). *Implementasi Bootstrap Sampling Pada Metode Ensemble Berdasarkan Metode Naïve Bayes, Random Forest Dan Regresi Logistik Biner Untuk Prediksi Kejadian Hujan* (Doctoral dissertation, Sriwijaya University).
- Nisa, K., Resti, Y., & Kresnawati, E. S. (2023). *Pengklasifikasian Kanker Serviks Menggunakan Metode Fuzzy Naïve Bayes Dengan Bootstrap Sampling* (Doctoral dissertation, Sriwijaya University).
- Laia, M. L., & Setyawan, Y. (2020). Perbandingan hasil klasifikasi curah hujan menggunakan metode SVM dan NBC. *Jurnal Statistika Industri dan Komputasi*, 5(02), 51-61.

- Lionovan, D. A., Santoso, L. W., & Intan, R. (2017). Klasifikasi Topik dan Analisa Sentimen Terhadap Kuesioner Umpan Balik Universitas Menggunakan Metode Long Short-Term Memory. *Jurnal Infra*, 8(2), 1–6.
- Nababan, C. E., & Simamora, E. (2023). Bootstrap Estimation of Confidence Intervals of Multiple Regression Model Parameters in the Presence of Multicollinearity Using Principal Component Analysis. *Formosa Journal of Applied Sciences*, 2(1), 185-202.
- Pamuji, F. Y., & Ramadhan, V. P. (2021). Komparasi Algoritma Random Forest dan Decision Tree untuk Memprediksi Keberhasilan Immunotherapy. *Jurnal Teknologi dan Manajemen Informatika*, 7(1), 46-50.
- Purba, L.I., Arsi., Armus, R., et al. (2021). *Agroklimatologi*. Medan: Yayasan Kita Menulis.
- Putra, A. I., & Santika, R. R. (2020). Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. *Edumatic : Jurnal Pendidikan Informatika*, 4(1), 121–130.
- Putra, P. A. D., Purnawan, I. A., & Putri, D. P. S. (2018). Sistem Pakar Diagnosa Penyakit Mata dengan Fuzzy Logic dan Naïve Bayes. *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, 6(1), 35.
- Putri, Z., Sugiyarto, S., & Salafudin, S. (2021). Sentiment analysis using fuzzy naïve bayes classifier on covid-19. *Desimal: Jurnal Matematika*, 4(2), 193-202.
- Rachman, F. P. A. P., Goejantoro, R., & Hayati, M. N. (2018). Penentuan Jumlah Replikasi Bootstrap Menggunakan Metode Pretest Pada Independent Sampel T Test. *Eksponensial*, 9(1), 35-40.
- Rahman, M. F., Alamsah, D., Darmawidjadja, M. I., & Nurma, I. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN). *Jurnal Informatika*, 11(1), 36.
- Rahmawan, H. (2020). Penentuan Rekomendasi Pelatihan Pengembangan Diri Bagi Pegawai Negeri Sipil Menggunakan Algoritma C4.5 Dengan Principal Component Analysis Dan Diskritisasi. *Jurnal Tekno Kompak*, 14(1), 5.
- Ramadhani, N., & Fajarianto, N. (2020). Sistem Informasi Evaluasi Perkuliahan dengan Sentimen Analisis Menggunakan Naïve Bayes dan Smoothing Laplace. *Jurnal Sistem Informasi Bisnis*, 10(2), 228–234.

- Randy., Hasniati., & Musdar, I. A. (2018). Aplikasi Prediksi Kerusakan Smartphone Menggunakan Metode Naive Bayes Dan Laplace Smoothing. *Jtriste*, 5(2), 8-16.
- Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*, 7(2), 156.
- Saputra, I., & Kristiyanti, D. A. (2021, Oktober 13). *Application of Data Mining for Rainfall Prediction Classification in Australia with Decision Tree Algorithm and C5. 0 Algorithm*. In Seminar Nasional Informatika (SEMNASIF), Yogyakarta.
- Setiawan, Y. E. (2020). Rekrutmen Guru Menggunakan Logika Fuzzy Tahani Supporting System of Decision Making of Teacher Recruitment Using Tahani Fuzzy Logic. *BAREKENG : Jurnal Ilmu Matematika Dan Terapan*, 14(2), 253–266.
- Sholihah, U. (2022). Penerapan Logika Fuzzy Tsukamoto untuk Memprediksi Curah Hujan di Kabupaten Kotawaringin Timur. *EJECTS: E-Journal Computer, Technology and ...*, 02(01), 31–39.
- Sunyoto, T., Setiabudi, D. H., & Tjondrowiguno, A. N. (2022). Penerapan Ensemble Learning Menggunakan Metode Support Vector Machine, Naïve Bayes Classifier, dan Valence Aware Dictionary for Sentiment Reasoning untuk Meningkatkan Akurasi Sentiment Analysis pada Review Aplikasi Google Play. *Jurnal Infra*, 10(2), 212–218.
- TP, B. P., & Sari, R. D. I. (2017). Penerapan Data Mining Untuk Prakiraan Cuaca Di Kota Malang Menggunakan Algoritma Iterative Dichotomiser Tree (Id3). *Joutica: Journal of Informatic Unisla*, 2(2).
- Wahono, H., & Riana, D. (2020). Prediksi Calon Pendorong Darah Potensial Dengan Algoritma Naïve Bayes, K-Nearest Neighbors dan Decision Tree C4.5. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 7.
- Winarno, G.D., Harianto, S.P., & Santoso, R. (2019). *Klimatologi Pertanian*. Bandar Lampung: Pusaka Media.