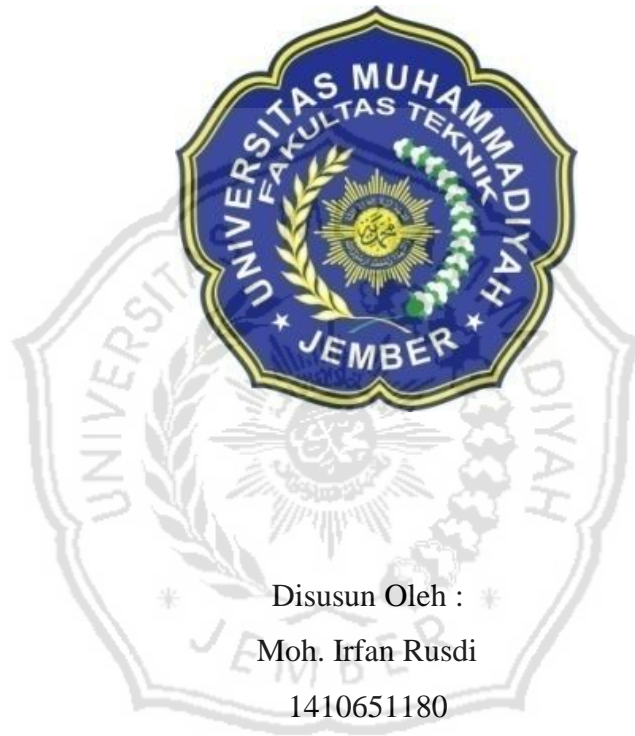


**TUGAS AKHIR**

**ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA  
NAÏVE BAYES UNTUK MENGLASIFIKASIKAN REKOMENDASI  
PENJUALAN ONLINE RETAIL**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH JEMBER  
2019**

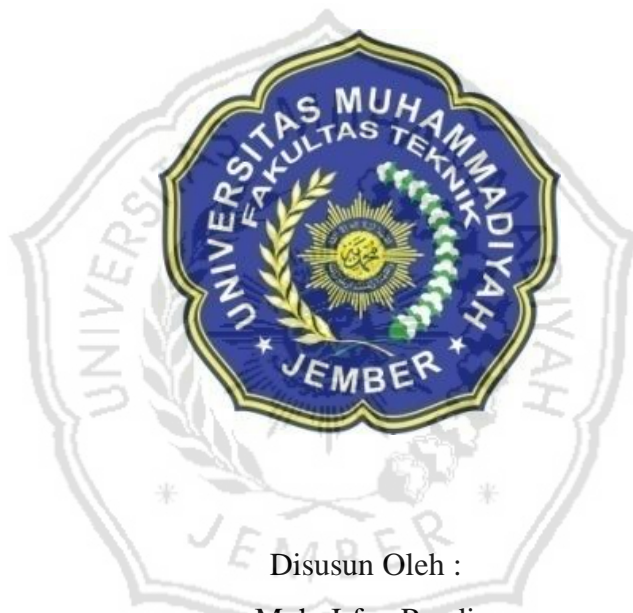
## **TUGAS AKHIR**

### **ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASIKAN REKOMENDASI PENJUALAN ONLINE RETAIL**

Sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom.)

Program Studi Teknik Informatika

Universitas Muhammadiyah Jember



Disusun Oleh :

Moh. Irfan Rusdi

1410651180

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH JEMBER**

**2019**

**HALAMAN PENGESAHAN**

**ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA  
NAÏVE BAYES UNTUK MENGLASIFIKASIKAN REKOMENDASI  
PENJUALAN ONLINE RETAIL**

**Mohammad Irfan Rusdi**

**1410651180**

Telah mempertanggung jawabkan Laporan Tugas Akhir pada Sidang Tugas Akhir  
tanggal 26 April 2019 sebagai salah satu syarat kelulusan dan mendapatkan

Gelar Sarjana Komputer (S.Kom)

Di

Universitas Muhammadiyah Jember

Disetujui oleh,

**Pembimbing I**

**Pembimbing II**

**Triawan Adi Cahyanto, M.Kom**

**Reni Umilasari, S.Pd, M.Si**

**NPK. 12 03 719**

**NPK. 19910728 1 1703813**

**Penguji I**

**Penguji II**

**Agung Nilogiri, S.T, M.Kom**

**Ilham Saifudin, S.Pd, M.Si**

**NIP. 1977 0330 2005 011002**

**NPK. 19891031 1 17 03811**

Mengesahkan,

**Dekan Fakultas Teknik**

Mengetahui,

**Ketua Program Studi Teknik**

**Informatika**

**Ir. Suhartinah, MT**

**NPK. 95 05 246**

**Yeni Dwi Rahayu, S.ST, M.Kom**

**NPK. 11 03 590**

## PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Moh. Irfan Rusdi

NIM : 1410651180

Institusi : Program Studi Teknik Informatika, Fakultas Teknik,  
Universitas Muhammadiyah Jember

Menyatakan bahwa tugas akhir yang berjudul “**ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASIKAN REKOMENDASI PENJUALAN ONLINE RETAIL**”, Bukan merupakan tugas akhir orang lain sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini di buat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar, penulis bersedia mendapatkan sanksi dari akademik.

Jember, 26 April 2019

**Mohammad Irfan Rusdi**

**1410651180**

## MOTTO

*“Bersabarlah kamu dan kuatkanlah kesabaranmu dan tetaplah bersiap-siaga dan bertaqwalah kepada Allah supaya kamu menang.”*

( QS. Al-Imraan : 200 )

*“Jika engkau ingin melihat indahnyafajar, maka engkau harus melalui gelapnya malam.”*

( Habib Syech bin Abdul Qadir Assegaf )



## UNGKAPAN TERIMA KASIH

Skripsi ini bukti dari terselesainya materi-materi mata kuliah yang telah di tempuh pada jenjang S1 Universitas Muhammadiyah Jember. Atas segala upaya, bimbingan dan arahan dari semua pihak. Penulis sangat berterima kasih yang sebesar-besarnya kepada :

1. Allah SWT beserta Nabi Muhammad SAW yang telah memberikan ridlo, kesehatan dan kelancaran atas terselesainya Tugas Akhir ini.
2. Kedua orang tua saya, Bapak Kuswandi dan Alm. Ibu Suwanti yang senantiasa mendoakan saya dengan penuh kasih sayang serta dukungan, support, restunya. Semoga Allah SWT memberikan kelapangan pintu rezeki terhadap beliau.
3. Kakak perempuan saya tercinta, Siti Qurrotul ‘Aini terimakasih telah menjadi ibu kedua saya serta support dan kasih sayangnya. Seseorang yang menginspirasi saya atas semangatnya.
4. Orang yang tersayang, Lum’atul ‘Aisyiyah, S.Pd terimakasih atas semua dukungan, support, waktu, tenaga, dan doa. Partner terbaik dalam hidupku.
5. Sahabat terbaik saya, Agil, Bang All, Deni, Rizki, Anas, Hendy dan Jen terimakasih telah menjadi tempat untuk bertukar pikiran.
6. Semua jajaran dosen Program Studi Teknik Informatika terimakasih atas ilmu yang tak akan putus sampai tua nanti.
7. Semua teman-teman angkatan 2014 Program Studi Teknik Informatika, terimakasih dukungan dan solidaritas tanpa batas.

## KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillah Puji syukur kehadirat Allah SWT yang telah memberi kekuatan kepada saya dalam menyelesaikan skripsi ini yang berjudul : “ANALISIS PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASIKAN REKOMENDASI PENJUALAN ONLINE RETAIL” yang dapat di selesaikan dengan segala kelebihan dan tak lepas dari kekurangan yang terdapat di dalamnya.

Sholawat serta salam semoga senantiasa tercurah kepada baginda Rosulullah Muhammad SAW. Serta keluarga beliau dan para sahabat hingga pengikutnya sampai akhir zaman, dan orang-orang yang senantiasa istiqomah menegakkan kebenaran dan menebar kebaikan di bumi hanya untuk Allah SWT.

Dengan segala kerendahan hati, penulis mohon maaf atas kekurangan yang masih ada dalam penulisan skripsi ini. Semoga hasil dari Tugas Akhir ini dapat mempermudah bagi pihak perusahaan dan pengembang agar berkembang lebih baik kedepannya. Semoga bermanfaat bagi setiap insan yang mempergunakannya untuk kebaikan di jalan Allah SWT.

Jember, 26 April 2019

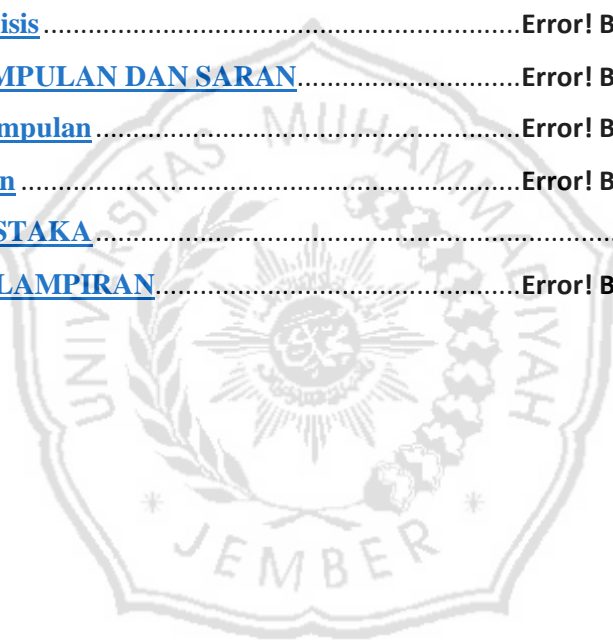
Penulis

## DAFTAR ISI

<u>HALAMAN PENGESAHAN</u> .....	ii
<u>PERNYATAAN KEASLIAN</u> .....	iii
<u>MOTTO</u> .....	iv
<u>UNGKAPAN TERIMA KASIH</u> .....	v
<u>KATA PENGANTAR</u> .....	vi
<u>DAFTAR ISI</u> .....	vii
<u>DAFTAR TABEL</u> .....	ix
<u>DAFTAR GAMBAR</u> .....	xi
<u>DAFTAR LAMPIRAN</u> .....	12
<u>ABSTRAK</u> .....	Error! Bookmark not defined.
<u>ABSTRACT</u> .....	Error! Bookmark not defined.
<u>BAB I PENDAHULUAN</u> .....	Error! Bookmark not defined.
<u>1.1. Latar Belakang</u> .....	Error! Bookmark not defined.
<u>1.2. Rumusan Masalah</u> .....	Error! Bookmark not defined.
<u>1.3. Batasan Masalah</u> .....	Error! Bookmark not defined.
<u>1.4. Tujuan Penelitian</u> .....	Error! Bookmark not defined.
<u>1.5. Manfaat Penelitian</u> .....	Error! Bookmark not defined.
<u>BAB II TINJAUAN PUSTAKA</u> .....	Error! Bookmark not defined.
<u>2.1 Online Retail</u> .....	Error! Bookmark not defined.
<u>2.2 Data Mining</u> .....	Error! Bookmark not defined.
<u>2.2.1 K-fold Cross Validation</u> .....	Error! Bookmark not defined.
<u>2.3 Klasifikasi</u> .....	Error! Bookmark not defined.
<u>2.4 Confusion Matrix</u> .....	Error! Bookmark not defined.
<u>2.4.1 Akurasi</u> .....	Error! Bookmark not defined.
<u>2.4.2 Presisi</u> .....	Error! Bookmark not defined.
<u>2.5 Decision Tree</u> .....	Error! Bookmark not defined.
<u>2.5.1 Algoritma C4.5</u> .....	Error! Bookmark not defined.
<u>2.6 Naïve Bayes</u> .....	Error! Bookmark not defined.
<u>2.7 Penelitian Terdahulu</u> .....	Error! Bookmark not defined.
<u>BAB III METODE PENELITIAN</u> .....	Error! Bookmark not defined.
<u>3.1 Studi Literatur</u> .....	Error! Bookmark not defined.
<u>3.2 Pengumpulan Data</u> .....	Error! Bookmark not defined.



<a href="#"><u>3.3</u></a>	<a href="#"><u>Preprocessing Data</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>3.4</u></a>	<a href="#"><u>Implementasi</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>3.4.1.</u></a>	<a href="#"><u>Perhitungan Data <i>Decission Tree</i> C4.5</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>3.4.2.</u></a>	<a href="#"><u>Perhitungan Data <i>Naïve Bayes</i></u></a> .....	Error! Bookmark not defined.
<a href="#"><u>3.5</u></a>	<a href="#"><u>Hasil</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>BAB IV HASIL DAN PEMBAHASAN</u></a> .....		Error! Bookmark not defined.
<a href="#"><u>4.1.</u></a>	<a href="#"><u>Implementasi Pada RapidMiner Studio</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>4.1.1.</u></a>	<a href="#"><u>Implementasi Algoritma C4.5</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>4.1.2.</u></a>	<a href="#"><u>Implementasi Algoritma <i>Naïve Bayes</i></u></a> .....	Error! Bookmark not defined.
<a href="#"><u>4.2.</u></a>	<a href="#"><u>Pengujian Algoritma C4.5</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>4.3.</u></a>	<a href="#"><u>Pengujian Algoritma <i>Naïve Bayes</i></u></a> .....	Error! Bookmark not defined.
<a href="#"><u>4.4.</u></a>	<a href="#"><u>Analisis</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>BAB V KESIMPULAN DAN SARAN</u></a> .....		Error! Bookmark not defined.
<a href="#"><u>5.1.</u></a>	<a href="#"><u>Kesimpulan</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>5.2.</u></a>	<a href="#"><u>Saran</u></a> .....	Error! Bookmark not defined.
<a href="#"><u>DAFTAR PUSTAKA</u></a> .....		13
<a href="#"><u>LAMPIRAN-LAMPIRAN</u></a> .....		Error! Bookmark not defined.



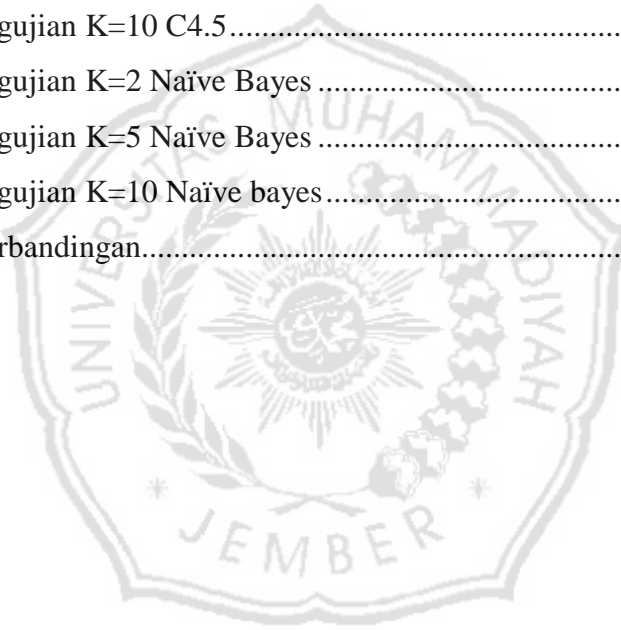
## DAFTAR TABEL

Tabel	Halaman
2.1 <i>Confusion Matrix</i> .....	9
2.2 Penelitian Terdahulu .....	14
3.1 Atribut data penjualan produk baju .....	19
3.2 Dataset Sampel .....	20
3.3 Jumlah kasus .....	20
3.4 Hasil perhitungan <i>entropy</i> .....	20
3.5 Hasil perhitungan <i>Gain</i> .....	23
3.6 <i>Entropy</i> dan <i>Gain</i> iterasi <i>Novelty</i> .....	25
3.7 <i>Entropy</i> dan <i>Gain</i> iterasi <i>Casual</i> .....	26
3.8 Dataset Sampel .....	28
3.9 Kelas $PC_1$ dan $PC_2$ .....	28
3.10 Perhitungan nilai Prior .....	29
3.11 Contoh data yang dihitung .....	31
3.12 Hasil klasifikasi <i>Naïve Bayes</i> .....	31
3.13 Rancangan Analisa C4.5 dan <i>Naïve Bayes</i> .....	32
4.1 Hasil pengujian C4.5 .....	39
4.2 <i>Confusion Matrix</i> $K=2$ .....	39
4.3 Hasil perhitungan per-skenario $K=2$ .....	39
4.4 <i>Confusion Matrix</i> $K=5$ .....	39
4.5 Hasil perhitungan per-skenario $K=5$ .....	40
4.6 <i>Confusion Matrix</i> $K=10$ .....	40
4.7 Hasil perhitungan per-skenario $K=10$ .....	40
4.8 <i>Confusion Matrix</i> $K=2$ .....	41
4.9 Hasil perhitungan per-skenario $K=2$ .....	41
4.10 <i>Confusion Matrix</i> $K=5$ .....	41
4.11 Hasil perhitungan per-skenario $K=5$ .....	41
4.12 <i>Confusion Matrix</i> $K=10$ .....	42
4.13 Hasil perhitungan per-skenario $K=10$ .....	42
4.14 <i>Confusion Matrix</i> $K=2$ .....	42

4.15 Hasil perhitungan per-skenario K=2 .....	43
4.16 Confusion Matrix K=5 .....	43
4.17 Hasil perhitungan per-skenario K=5 .....	43
4.18 Confusion Matrix K=10 .....	44
4.19 Hasil perhitungan per-skenario K=10 .....	44
4.20 Hasil pengujian Naïve Bayes .....	44
4.21 Confusion Matrix K=2 .....	45
4.22 Hasil perhitungan per-skenario K=2 .....	45
4.23 Confusion Matrix K=5 .....	45
4.24 Hasil perhitungan per-skenario K=5 .....	45
4.25 Confusion Matrix K=10 .....	46
4.26 Hasil perhitungan per-skenario K=10 .....	46
4.27 Confusion Matrix K=2 .....	46
4.28 Hasil perhitungan per-skenario K=2 .....	47
4.29 Confusion Matrix K=5 .....	47
4.30 Hasil perhitungan per-skenario K=5 .....	47
4.31 Confusion Matrix K=10 .....	47
4.32 Hasil perhitungan per-skenario K=10 .....	48
4.33 Confusion Matrix K=2 .....	48
4.34 Hasil perhitungan per-skenario K=2 .....	48
4.35 Confusion Matrix K=5 .....	49
4.36 Hasil perhitungan per-skenario K=2 .....	49
4.37 Confusion Matrix K=10 .....	49
4.38 Hasil perhitungan per-skenario K=10 .....	49
4.39 Hasil tertinggi pada setiap skenario uji C4.5 .....	50
4.40 Hasil tertinggi pada setiap skenario uji Naïve Bayes .....	51
4.41 Hasil tertinggi C4.5 .....	51
4.42 Hasil tertinggi Naïve Bayes .....	51
4.43 Perbandingan hasil tertinggi pada algoritma C4.5 dan Naïve Bayes ...	52

## DAFTAR GAMBAR

Gambar	Halaman
2.1 Tahapan <i>Data Mining</i> .....	7
2.2 Proses <i>5-fold Cross Validation</i> .....	7
2.3 Proses pekerjaan <i>Klasifikasi</i> .....	8
3.1 Alur Penelitian .....	17
3.2 Flowchart .....	19
3.4 Pohon Keputusan .....	27
4.1 Hasil pengujian K=2 C4.5.....	35
4.2 Hasil pengujian K=5 C4.5.....	36
4.3 Hasil pengujian K=10 C4.5.....	36
4.4 Hasil pengujian K=2 Naïve Bayes .....	37
4.5 Hasil pengujian K=5 Naïve Bayes .....	38
4.6 Hasil pengujian K=10 Naïve bayes.....	38
4.7 Grafik perbandingan.....	52



## DAFTAR LAMPIRAN

Lampiran	Halaman
1 Dataset .....	59
2 Dataset uji .....	73
3 Perhitungan $PC_1$ dan $PC_2$ pada algoritma Naïve Bayes .....	106
4 Proses pengujian .....	108
5 Lembar Revisi .....	112
6 Biodata .....	113



## DAFTAR PUSTAKA

- Fatmawati. (2016). Perbandingan Algoritma Klasifikasi Data Mining Model C4.5 dan Naive Bayes untuk prediksi penyakit diabetes. *Jurnal Techno Nusa Mandiri Vol. XIII, No. 1*.
- Fitrianah. Arifin, M. F., Devi. (2017). Penerapan Algoritma Klasifikasi C4.5 Dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada, *Jurnal Telekomunikasi dan Komputer*, vol.8, no.2, 2018
- Kusrini Dan Emha. (2009). *Algoritma Data Mining*. Yogyakarta: C.V Andi.
- Larose. (2014). *Discovering Knowledge In Data*. New Jersey: John Wiley & Sons.
- Muflikhah, L., Ratnawati, D. E., Putri, R. R. M. (2018). *Data Mining*. Malang: Ub Press.
- Muljono., Artanti, D. P., et. al. (2018). Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naive Bayes. *Konferensi Nasional Sistem Informasi, STMIK Atma Luhur Pangkalpinang*, 8 – 9 Maret 2018.
- Pattipeilohy, W. F., Wibowo, A., Utari, D. R. (2017). Pemodelan Dan Prototipe Sistem Informasi Untuk Prediksi Pembaharuan Polis Asuransi Mobil Menggunakan Algoritma C.45. *Prosiding SNATIF*. 783.
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi menggunakan MATLAB*. Yogyakarta: C.V Andi OFFSET.
- Rahmawati, E. (2015). Analisa Komparasi Algoritma Naive Bayes Dan C4.5 Untuk Prediksi Penyakit Liver. *Jurnal Techno Nusa Mandiri*, Vol. XII No. 2, September 2015.
- Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4.5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus : Kspss / Bmt Al-Fadhila). *Magister Teknik Informatika Institut Informatika dan Bisnis Darmajaya*, Vol. 02 No. 01 Mei 2016.
- Septiani, W. D. (2017). Komparasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis. *Jurnal Pilar Nusa Mandiri*. Volume 13 No.1, Maret 2017.

Sulianta, F., Dominikus, J. (2010). *Data Mining Meramalkan Bisnis Perusahaan*. Jakarta: PT. Elex Media Komputindo.

Susanto, E. D. (2015). Penerapan Data Mining Classification Untuk Prediksi Perilaku

Pola Pembelian Terhadap Waktu Transaksi Menggunakan Metode Naïve Bayes. *Konferensi Nasional Sistem & Informatika*. STMIK STIKOM Bali, 9 – 10 Oktober 2015.

Widiyanto, I., Prasilowati. S. L. (2015). Perilaku Pembelian Melalui Internet, *Jurnal Manajemen Dan Kewirausahaan*, Vol.17, No. 2, September 2015.

Wardani, N. W., Dantes, G. R., Indrawan, G. (2018). Prediksi Customer Churn Dengan Algoritma Decision Tree C4.5 Berdasarkan Segmentasi Pelanggan Pada Perusahaan Retail. *Jurnal Resistor*, Vol. 1 No 1 – April 2018.

