

***DATA MINING MENGGUNAKAN ALGORITMA K-NEAREST  
NEIGHBOR UNTUK ANALISA PERILAKU PELANGGAN PADA  
SISTEM LAYANAN PENGGALANGAN DANA ONLINE***

**Tesis  
untuk memenuhi sebagian persyaratan  
mencapai derajat Sarjana S-2  
Program Studi Magister Sistem Informasi**



**Chashif Syadzali  
30000416410032**

**SEKOLAH PASCASARJANA**

**SEKOLAH PASCASARJANA  
UNIVERSITAS DIPONEGORO  
SEMARANG**

**2020**

**HALAMAN PENGESAHAN**

**TESIS**

***DATA MINING MENGGUNAKAN ALGORITMA K-NEAREST  
NEIGHBOR UNTUK ANALISA PERILAKU PELANGGAN PADA  
SISTEM LAYANAN PENGGALANGAN DANA ONLINE***

**Oleh:  
Chashif Syadzali  
30000416410032**

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 6 Agustus 2020 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Pembimbing I

Semarang, Agustus 2020  
Mengetahui,  
Penguji I

Dr. Suryono, S.Si, M.Si  
NIP. 197306301998021001

Dr. Rahmat Gernowo, M.Si  
NIP. 196511231994031000

Pembimbing II

Penguji II

Jatmiko Endro Suseno, M.Si., Ph.D  
NIP. 197211211998021001

V. Gunawan S.K., M.Si., Ph.D  
NIP. 197105221997021000

Mengetahui,  
Ketua Program Studi  
Magister Sistem Informasi UNDIP

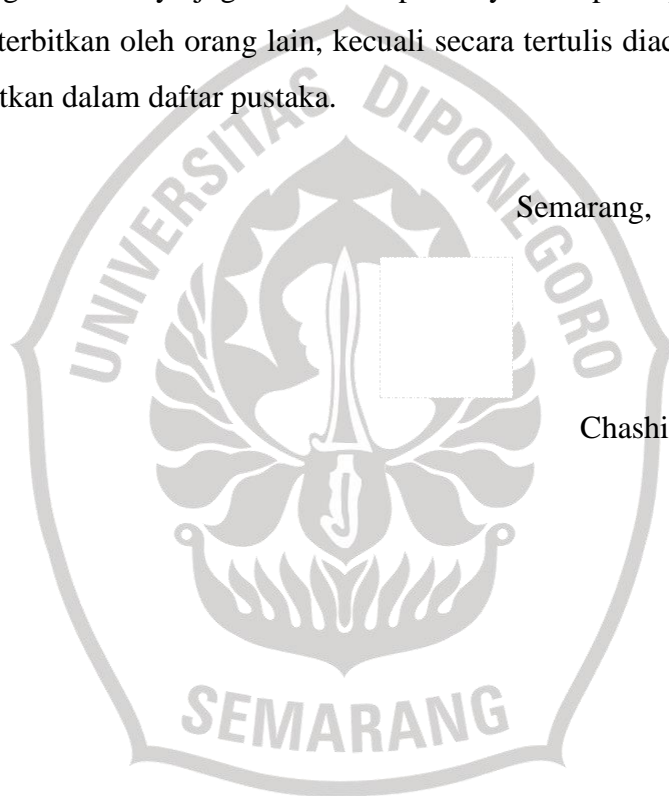
Dr. Suryono, S.Si, M.Si  
NIP. 197306301998021001

## HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa dalam tesis ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Semarang, Agustus 2020

Chashif Syadzali



**PERNYATAAN PERSETUJUAN  
PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini :

Nama : Chashif Syadzali  
NIM : 30000416410032  
Program Studi : Magister Sistem Informasi  
Program : Pascasarjana  
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif atas karya ilmiah saya yang berjudul :

***DATA MINING MENGGUNAKAN ALGORITMA K-NEAREST  
NEIGHBOR UNTUK ANALISA PERILAKU PELANGGAN PADA  
SISTEM LAYANAN PENGGALANGAN DANA ONLINE***

beserta perangkat yang ada. Dengan Hak bebas Royalti Noneksklusif ini Magister Sistem Informasi Pascasarjana Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Dibuat di : Semarang

Pada tanggal : 6 Agustus 2018

Yang menyatakan

Chashif Syadzali  
NIM. 30000416410032

## KATA PENGANTAR

*Alhamdulillah* puji syukur atas rahmat Allah SWT, Segala puji bagi Allah atas berkat, rahmat, taufik dan hidayah-Nya, penyusunan Tesis ini dapat diselesaikan karena didukung dari berbagai pihak yang telah membantu. Oleh karena itu penulis menyampaikan banyak terima kasih kepada :

1. Bapak Prof. Dr. Suryono, S.Si, MSi selaku pembimbing I
2. Bapak Jatmiko Endro Suseno, M.Si., Ph.D selaku pembimbing II
3. Bapak Prof. Dr. Suryono, S.Si, MSi selaku Ketua Program Studi Magister Sistem Informasi Universitas Diponegoro Semarang

Dengan keterbatasan pengalaman, ilmu maupun pustaka yang ditinjau, penulis menyadari bahwa tesis ini masih banyak kekurangan dan membutuhkan pengembangan lanjut agar lebih bermanfaat. Oleh sebab itu, penulis sangat mengharapkan kritik dan saran agar tesis ini lebih sempurna serta sebagai masukan bagi penulis untuk penelitian dan penulisan karya ilmiah di masa yang akan datang. Akhir kata, penulis berharap tesis ini memberikan manfaat bagi kita semua terutama untuk pengembangan ilmu pengetahuan

Semarang, Agustus 2020

Penulis

SEKOLAH PASCASARJANA

## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN .....	iii
PERNYATAAN PERSETUJUAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR .....	viii
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
ABSTRAK .....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Tujuan Penelitian.....	4
1.3. Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Tinjauan Pustaka.....	5
2.2. Dasar Teori .....	7
2.2.1. <i>Customer Lifecycle</i> .....	7
2.2.2. <i>Data Mining</i> .....	9
2.2.3. <i>K-Nearest Neighbor</i> .....	10
2.2.4. <i>Business Intelligence</i> .....	11
2.2.5. <i>Crowdfunding</i> .....	11
BAB III METODE PENELITIAN.....	13
3.1. Bahan dan Alat Penelitian.....	13
3.2. Prosedur Penelitian .....	13

3.3. Perancangan Sistem.....	15
3.3.1. Kerangka Sistem Informasi .....	15
3.3.2. Perancangan Basis Data.....	16
3.3.3. Perancangan Proses.....	19
3.3.4. Perancangan Antarmuka.....	24
3.3.5. Pengkodean Program .....	28
BAB IV HASIL DAN PEMBAHASAN PENELITIAN.....	32
4.1. Hasil Penelitian.....	32
4.1.1. Implementasi Sistem.....	35
4.1.2. Perhitungan <i>K-Nearest Neighbor</i> .....	43
4.2. Pengujian Sistem .....	45
4.2.1. Pengujian Akurasi.....	45
4.2.2. Pengujian Sistem dengan <i>Blackbox Testing</i> .....	46
4.2.3. Pengujian Modul.....	48
BAB V PENUTUP.....	51
5.1. Kesimpulan.....	51
5.2. Saran .....	51
DAFTAR PUSTAKA .....	52
LAMPIRAN.....	54



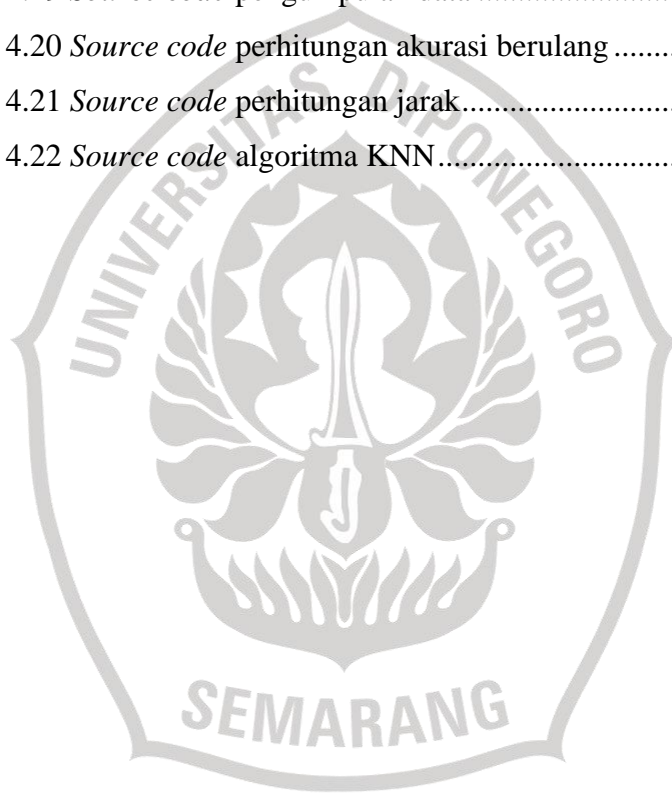


## DAFTAR GAMBAR

Gambar 3.1 Pengembangan sistem model <i>waterfall</i> .....	14
Gambar 3.2 Kerangka Sistem Informasi .....	16
Gambar 3.3 Diagram konteks .....	19
Gambar 3.4 Diagram level 1 .....	20
Gambar 3.5 DFD level 2 proses 1 .....	22
Gambar 3.6 DFD level 2 proses 2 .....	22
Gambar 3.7 DFD level 2 proses 3 .....	23
Gambar 3.8 DFD level 2 proses 4 .....	23
Gambar 3.9 <i>Form Login</i> .....	24
Gambar 3.10 Rancangan halaman data donasi .....	25
Gambar 3.11 Rancangan halaman data pengguna .....	25
Gambar 3.12 Rancangan halaman data kampanye .....	26
Gambar 3.13 Rancangan halaman klasifikasi pola perilaku pengguna .....	27
Gambar 3.14 Rancangan halaman analisa pola perilaku pengguna .....	28
Gambar 3.15 Perintah <i>SQL select</i> untuk mendapatkan data pengguna .....	28
Gambar 4.1 Data Donasi .....	33
Gambar 4.2 Data Pengguna .....	33
Gambar 4.3 Data Kampanye .....	34
Gambar 4.4 Hasil Klasifikasi menggunakan <i>K-Nearest Neighbor</i> .....	34
Gambar 4.5 <i>Input</i> , proses dan <i>output</i> sistem .....	35
Gambar 4.6 Pemilihan data .....	36
Gambar 4.7 Pengumpulan data latih untuk klasifikasi .....	36
Gambar 4.8 Melakukan akurasi perhitungan KNN .....	37
Gambar 4.9 Hasil klasifikasi pengguna .....	38
Gambar 4.10 Percobaan dengan nilai $K=3$ .....	38
Gambar 4.11 Hasil klasifikasi pengguna menggunakan KNN .....	39
Gambar 4.12 Diagram persentase klasifikasi pengguna .....	40
Gambar 4.13 Performa KNN .....	41

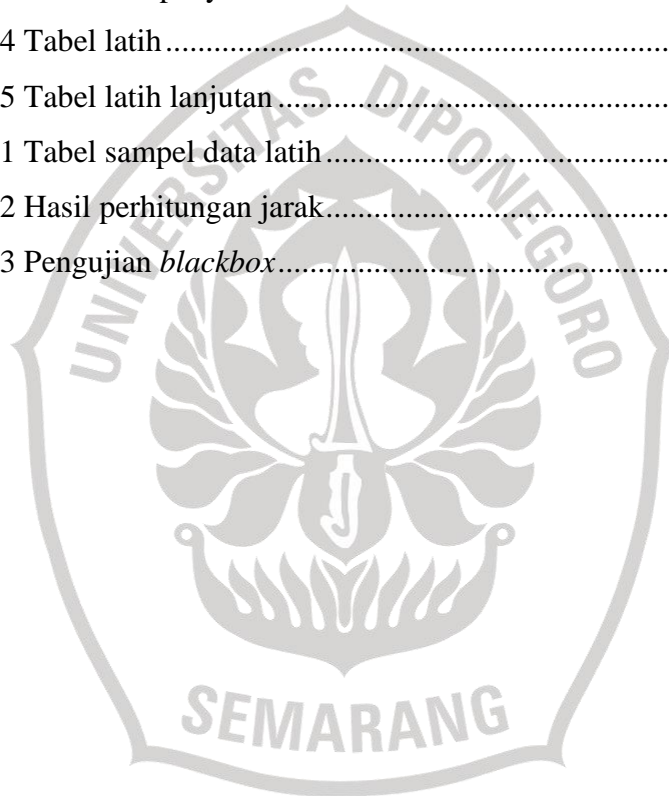


Gambar 4.14 Grafik donasi.....	41
Gambar 4.15 <i>Summary</i> pengguna.....	42
Gambar 4.16 <i>Latest member</i> .....	42
Gambar 4.17 <i>Latest donation</i> .....	43
Gambar 4.18 Perhitungan akurasi berulang.....	45
Gambar 4.19 <i>Source code</i> pengumpulan data .....	48
Gambar 4.20 <i>Source code</i> perhitungan akurasi berulang .....	49
Gambar 4.21 <i>Source code</i> perhitungan jarak.....	49
Gambar 4.22 <i>Source code</i> algoritma KNN.....	50



## DAFTAR TABEL

Tabel 3.1 Tabel pengguna.....	17
Tabel 3.2 Tabel donasi.....	17
Tabel 3.3 Tabel kampanye.....	18
Tabel 3.4 Tabel latihan.....	18
Tabel 3.5 Tabel latihan lanjutan.....	19
Tabel 4.1 Tabel sampel data latihan.....	43
Tabel 4.2 Hasil perhitungan jarak.....	44
Tabel 4.3 Pengujian <i>blackbox</i> .....	47



## DAFTAR LAMPIRAN

Lampiran 1 Sampel data latih .....	55
Lampiran 2 Data hasil perhitungan jarak.....	56
Lampiran 3 Data pengguna.....	57



## ABSTRAK

Klasifikasi pola perilaku pengguna dapat bermanfaat untuk membantu perusahaan dalam melakukan analisa *business intelligence*. Teknik *data mining* dapat melakukan klasifikasi pola perilaku pengguna menggunakan algoritma *K-Nearest Neighbor* berdasarkan siklus hidup pelanggan yang terdiri dari *prospect*, *responder*, *active* dan *former*. Data yang digunakan untuk melakukan klasifikasi meliputi usia, jenis kelamin, jumlah donasi, retensi donasi dan jumlah kunjungan pengguna. Hasil perhitungan dari 2.114 data menghasilkan klasifikasi masing-masing kategori pengguna yaitu *active* sebesar 1,18%, *prospect* sebesar 8,99%, *responder* sebesar 4,26% dan *former* sebesar 85,57%. Akurasi sistem menggunakan rentang  $K$  dari  $K=1$  sampai  $K=20$  menghasilkan akurasi tertinggi pada nilai  $K=4$  dengan akurasi sebesar 94,37%. Hasil dari pembelajaran data latih yang menghasilkan klasifikasi pola perilaku pengguna dapat digunakan sebagai analisa *Business Intelligence* yang bermanfaat untuk perusahaan dalam menentukan strategi bisnis dengan mengetahui target pasar optimal.

Kata Kunci: Klasifikasi, *data mining*, *K-Nearest Neighbor*, *Business Intelligence*, segmentasi pengguna, siklus hidup pelanggan, *Customer Relationship Management*.



SEKOLAH PASCASARJANA

## ***ABSTRACT***

Customer behavior classification could be useful to assist companies in conducting business intelligence analysis. Data mining techniques classified customer behavior using the K-Nearest Neighbor algorithm based on the customer's life cycle consisting of prospect, responder, active and former. Data used to classification included age, gender, number of donations, donation retention and number of user visits. The calculation resulted from 2,114 data in the classification of each customer's category are namely active by 1.18%, prospect by 8.99%, responder by 4.26% and former by 85.57%. System accuracy using a range of K from K = 1 to K = 20 produces that the highest accuracy was 94.3731% at a value of K = 4. The results of the training data that produced a classification of user behavior was used as a Business Intelligence analysis for companies in determining business strategies by knowing the target of optimal market.

Keyword: Classification, data mining, K-Nearest Neighbor, Business intelligence, user segmentation, customer life cycle, Customer Relationship Management.

