

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR BERBASIS COSINE SIMILARITY UNTUK KLASIFIKASI TEKS KELUHAN MASYARAKAT

Lutfiyah Furaida
(13.04.111.000.81)

Pembimbing I : Achmad Jauhari, S.T., M.Kom.
Pembimbing II : Yoga Dwitya Pramudita, SKom., M.Cs.

ABSTRAK

Kota Bangkalan merupakan salah satu Kabupaten yang ada di Pulau Madura. Kota yang sudah cukup berkembang dengan berbagai kemajuan teknologi dan informasi saat ini, akan tetapi kemajuan tersebut tidak menyeluruh. Seperti halnya pelayanan publik dimana ketika masyarakat ingin menyampaikan keluhan atau ingin tahu informasi terbaru, masyarakat masih harus datang ke Dinas terkait. Diperlukan klasifikasi teks untuk menentukan kelas kategori pada keluhan. Penelitian ini menggunakan data keluhan dari dinas terkait di kabupaten bangkalan dan akan dibangun aplikasi keluhan masyarakat dengan mengimplementasikan algoritma *K-Nearest Neighbor* (KNN) berbasis *Cosine Similarity*. Parameter nilai *K* ditentukan dengan $K=4$ sampai $K=10$ sehingga menemukan nilai *K* optimal. Hasil penelitian menunjukkan bahwa *K* optimal dengan rata-rata akurasi 97,5% di peroleh ketika nilai $K=4$. Selain itu, dengan adanya aplikasi keluhan masyarakat ini dapat mempermudah pengguna dalam menyampaikan keluhan, karena keluhan dapat diterima oleh masyarakat lain yang belum mengetahui informasi keluhan yang ada disekitar kabupaten bangkalan dengan cepat dan mudah.

Kata kunci : *Klasifikasi, K-Nearest Neighbor (KNN).*

KATA PENGANTAR

Puji Syukur kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, serta hidayah-Nya, sehingga laporan tugas akhir ini yang berjudul **Penerapan Algoritma K-Nearest Neighbor berbasis Cosine Similarity untuk klasifikasi teks keluhan masyarakat** ini dapat terselesaikan dengan baik.

Dalam penyusunan tugas akhir ini penulis telah mendapat banyak bimbingan, saran, petunjuk serta dorongan dari berbagai pihak. Oleh karena itu, perkenankan penulis untuk mengucapkan terima kasih kepada :

1. Allah SWT, Sang Maha Pencipta yang telah memberikan rahmat, hidayah, anugerah, ridho, serta taufik-Nya kepada penulis sehingga penulis dapat menyelesaikan Tugas Akhir ini. Dan sholawat serta salam tetap terlimpahkan kepada junjungan Nabi besar kita Nabi Muhammad SAW.
2. Kedua orang tuaku yang selalu sabar dan mendukung setiap langkahku, adikku Sahda Naura Rizkiyah. Dan keluarga besar bani H.soetomo dan bani abu ali.
3. Dosen Pembimbing Bapak Achmad Jauhari, S.T., M.Kom. dan Bapak Yoga Dwitya Pramudita, SKom., M.Cs. yang telah membantu banyak penulis dalam menyelesaikan Tugas Akhir ini.
4. Seluruh Bapak Ibu Dosen Program Studi Teknik Informatika Universitas Trunojoyo Madura.
5. Seluruh teman-teman Program Studi Teknik Informatika Universitas Trunojoyo Madura khususnya angkatan 2013 (Increment).
6. Teman-teman Kos Az-Zahra 2, Kong Orange dan Mas om yang memberikan dukungan dan bantuan tenaga sepenuhnya.

Bangkalan, 24 Juli 2016

Penulis

DAFTAR ISI

Lembar Pengesahan	v
Halaman Pernyataan Orisinalitas	vii
Halaman Persembahan	ix
Abstrak.....	xi
Kata Pengantar	xiii
Daftar Isi	xv
Daftar Gambar	xvii
Daftar Tabel	xix
Daftar Lampiran.....	xx
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan dan Manfaat	3
1.4. Batasan Masalah	3
1.5. Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Penelitian Sebelumnya	5
2.2. Landasan Teori.....	8
2.2.1 <i>Smart City</i>	8
2.2.2 Proses TF-IDF	9
2.2.3 Metode K-Nearest Neighbor(KNN)	9
2.2.4 <i>Cosine Similarity</i>	10
2.2.5 <i>Confusion Matrix</i>	10
2.2.6 Basis Data	11
2.2.7 <i>MySQL</i>	12
2.3. Contoh Perhitungan Manual	12
BAB III ANALISA DAN PERANCANGAN SISTEM	19
3.1. Analisa Sistem	19
3.1.1. Gambaran Umum Sistem	19
3.1.2. Analisa Kebutuhan Fungsional	20
3.1.2. Analisa Kebutuhan Non-Fungsional	20
3.2. Desain Sistem.....	21
3.2.1. <i>Fowchart</i> Sistem Admin	22
3.2.2. <i>Fowchart</i> Sistem User.....	23
3.2.3. <i>Fowchart</i> Kalsifikasi Teks	24
3.2.4. <i>Use Case Diagram</i>	26
3.2.5. Arsitektur Sistem.....	29

3.2.6. <i>Activity Diagram</i>	30
3.2.7. <i>Sequence Diagram</i>	43
3.2.8. <i>User Interface</i>	56
3.3. Data Modeling	60
3.3.1. <i>Conceptual Data Modeling (CDM)</i>	61
3.3.2. <i>Physical Data Modeling (CDM)</i>	61
3.4. Struktur Tabel	63
3.5. Rancangan Skenario Uji Coba	68
BAB IV IMPLEMENTASI DAN PENGUJIAN	69
4.1. Lingkungan Uji Coba.....	69
4.2. Pembuatan Sistem	69
4.2.1. Pengaturan <i>Database</i>	69
4.2.2. Perhitungan KNN	70
4.3. Implementasi Program	72
4.3.1. Halaman Form Keluhan.....	72
4.3.2. Halaman Klasifikasi Keluhan.....	73
4.3.3. Halaman View Lokasi Keluhan.....	74
4.4. Uji Coba Sistem.....	74
4.4.1. Uji Fungsional.....	75
4.4.2. Uji Klasifikasi Sistem.....	79
4.4.3. Analisa Sistem.....	80
BAB V PENUTUP	83
5.1. Kesimpulan	83
5.2. Saran	83
DAFTAR PUSTAKA	85
LAMPIRAN.....	87
BIODATA.....	100

DAFTAR GAMBAR

Gambar 2.1	Enam jenis smart city menurut IBM.....	8
Gambar 3.1	<i>Flowchart</i> Sistem Admin.....	22
Gambar 3.2	<i>Flowchart</i> Sistem User	23
Gambar 3.3	<i>Flowchart</i> Klasifikasi Teks.....	24
Gambar 3.4	<i>Use Case Diagram</i>	26
Gambar 3.5	Arsitekktur Sistem	29
Gambar 3.6	<i>Activity</i> Login Admin	30
Gambar 3.7	<i>Activity</i> Input Kategori Admin.....	31
Gambar 3.8	<i>Activity</i> Edit Kategori Admin	32
Gambar 3.9	<i>Activity</i> Cari Kategori Admin	33
Gambar 3.10	<i>Activity</i> Hapus Kategori Admin.....	34
Gambar 3.11	<i>Activity</i> Input Data Training Admin	35
Gambar 3.12	<i>Activity</i> Hapus Data Training Admin.....	36
Gambar 3.13	<i>Activity</i> Cari Data Training Admin	37
Gambar 3.14	<i>Activity</i> Input Keluhan Admin	38
Gambar 3.15	<i>Activity</i> Cari Keluhan Admin.....	39
Gambar 3.16	<i>Activity</i> View Lokasi Keluhan Admin	40
Gambar 3.17	<i>Activity</i> Input Keluhan User.....	41
Gambar 3.18	<i>Activity</i> Cari Keluhan User	42
Gambar 3.19	<i>Activity</i> View Lokasi Keluhan User.....	43
Gambar 3.20	<i>Sequence</i> Login Admin	44
Gambar 3.21	<i>Sequence</i> Input Kategori Admin.....	45
Gambar 3.22	<i>Sequence</i> Edit Kategori Admin	46
Gambar 3.23	<i>Sequence</i> Cari Kategori Admin	47
Gambar 3.24	<i>Sequence</i> Hapus Kategori Admin.....	48
Gambar 3.25	<i>Sequence</i> Input Data Training Admin	49
Gambar 3.26	<i>Sequence</i> Hapus Data Training Admin.....	50
Gambar 3.27	<i>Sequence</i> Input Keluhan Admin	51
Gambar 3.28	<i>Sequence</i> Cari Keluhan Admin.....	52
Gambar 3.29	<i>Sequence</i> View Lokasi Keluhan Admin	53
Gambar 3.30	<i>Sequence</i> Input Keluhan User.....	54
Gambar 3.31	<i>Sequence</i> Cari Keluhan User	55
Gambar 3.32	<i>Sequence</i> View Lokasi Keluhan User.....	56
Gambar 3.33	Halaman <i>Login Admin</i>	56
Gambar 3.34	Halaman <i>Home Admin</i>	57
Gambar 3.35	Halaman <i>Home User</i>	58
Gambar 3.36	Halaman <i>View Lokasi</i>	59

Gambar 3.37	<i>Conceptual Data Modeling (CDM)</i>	61
Gambar 3.38	<i>Physical Data Modeling (CDM)</i>	62

DAFTAR TABEL

Tabel 2.1	Analisis penelitian sebelumnya	6
Tabel 2.2	<i>Confusion Matrix</i>	10
Tabel 2.3	Data Keluhan	13
Tabel 2.4	TF-IDF	14
Tabel 2.5	Perhitungan <i>Cosine</i>	17
Tabel 2.6	Pengurutan <i>Cosine</i>	17
Tabel 3.1	<i>User Requirment</i>	27
Tabel 3.2	Diskripsi <i>Use Case</i>	27
Tabel 3.3	T_user	63
Tabel 3.4	T_dok	63
Tabel 3.5	T_kategori	64
Tabel 3.6	T_keluhan	64
Tabel 3.7	T_unik	64
Tabel 3.8	T_Term	65
Tabel 3.9	T_Wdt	65
Tabel 3.10	T_Pjgdok	65
Tabel 3.11	T_unikkey	66
Tabel 3.12	T_termkey	66
Tabel 3.13	T_vektor	66
Tabel 3.14	T_katadasar	67
Tabel 3.15	T_stoplist	67
Tabel 3.16	T_hasil	67
Tabel 3.17	Skenario Uji Coba	68
Tabel 4.1	Kode database	70
Tabel 4.2	Kode Program <i>Preprocessing</i>	70
Tabel 4.3	Kode Program TF-IDF	70
Tabel 4.4	Kode Program <i>Cosine Similarity</i>	71
Tabel 4.5	Kode Program Perengkingan	71
Tabel 4.6	Kode Program Mayoritas	71
Tabel 4.7	Kode Program Hasil	72
Tabel 4.8	Uji fungsionalitas level admin	75
Tabel 4.9	Uji fungsionalitas level User	78
Tabel 4.10	<i>Hasil Confusion Matrix</i>	79
Tabel 4.11	Hasil Skenario Kedua	79

DAFTAR LAMPIRAN

Lampiran 1	Kuesioner Uji Klasifikasi	87
Lampiran 2	Sourcode KNN	91
Lampiran 3	Sourcode Get K	96
Lampiran 4	Hasil perhitungan manual.....	98