

MODEL DATA MINING PELANGGAN PAM JAYA DKI JAKARTA DENGAN TEKNIK *EMERGING PATTERN*

TESIS

Diajukan untuk memenuhi salah satu persyaratan
memperoleh gelar Magister Ilmu Komputer (M.KOM)



Oleh:

MOHAMMAD ICHLAS

1311600215

PROGRAM STUDI MAGISTER ILMU KOMPUTER (M.KOM)

PROGRAM PASCASARJANA

UNIVERSITAS BUDI LUHUR

JAKARTA

2016



**PROGRAM STUDI MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama Mahasiswa : Mohamad Ichlas

Nomor Induk Mahasiswa : 1311600215

Konsentrasi : Teknologi Sistem Informasi

Program Studi : Magister Ilmu Komputer

Judul Proposal Tesis : Model data mining pelanggan PAM JAYA DKI Jakarta dengan teknik *Emerging Pattern*

Menyatakan bahwa TESIS yang berjudul :

MODEL DATA MINING PELANGGAN PAM JAYA DKI JAKARTA DENGAN TEKNIK *EMERGING PATTERN*.

1. Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan hasil karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain.
2. Saya ijin untuk dikelola oleh Universitas Budi Luhur sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya bersedia menerima konsekuensi apapun sesuai dengan aturan apabila pernyataan ini tidak benar.

Jakarta, 25 Juni 2016

BIETERAI TEMPEL
6000



PROGRAM STUDI MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama Mahasiswa : Mohammad Ichlas
Nomor Induk Mahasiswa : 1311600215
Konsentrasi : Teknologi Sistem Informasi
Program Studi : Magister Ilmu Komputer
Judul Tesis : Model data mining pelanggan PAM JAYA DKI
Jakarta dengan teknik *Emerging Pattern*

Jakarta, 25 Juni 2016

Tim Penguji

Ketua,
Prof. Dr. Moedjiono, M.Sc

Tanda Tangan

Anggota,
Dr., Ir. Nazori Az, M.T

Pembimbing Utama,
Rusdah, M.Kom

Pembimbing Pendamping,
Spits Warnars Harco Leslie Hendric, Ph.D

Ketua Program Studi

(Dr. Mohammad Syafrullah, M.Kom, M.Sc)

ABSTRAK

Emerging Pattern merupakan salah satu metodologi pengelompokan data dengan pemolaan data yang sangat terstruktur. *Emerging Pattern* ditujukan untuk membantu *top management* di PAM DKI Jakarta untuk membuat keputusan. Permasalahan utama yang dihadapi oleh Perusahaan Air Minum (PAM) DKI Jakarta adalah lambannya layanan dalam memperluas cakupan jumlah pelanggannya. Ketika pelanggan baru ingin mengajukan pemasangan saluran air, pelanggan menunggu lama realisasinya. Hal ini disebabkan karena data-data lokasi pelanggan baru harus disesuaikan dengan data-data lokasi dari pipa saluran air utama membutuhkan waktu yang lama sehingga waktu tunggu untuk pemasangan saluran air sangat lama. Oleh karena itu, untuk menentukan operator mana yang berhak mendapat program, diperlukan suatu sistem yang dapat membantu untuk pengambilan keputusan atau bisa disebut sistem pendukung keputusan. Salah satu metode yang dapat digunakan untuk membuat sistem pendukung keputusan yaitu Metode Simple Additive Weighting (SAW). Metode ini dipilih karena dinilai mampu menyeleksi alternatif terbaik dari sejumlah alternatif, dalam hal ini alternatif yang dimaksud yaitu calon operator yang berhak menerima program didasarkan atas kriteria-kriteria tertentu. Penilaian dilakukan dengan cara menentukan pengelompokan atau pengclusteran dari data pelanggan, kemudian menghitung indeks XB (Xie-Beni) pada tiap - tiap cluster yang telah terbentuk sehingga dari perhitungan tersebut diketahui cluster mana yang paling baik yang dapat dijadikan alternatif untuk pengambilan keputusan, setelah itu dilakukan proses perangkingan dari setiap cluster yang akan menentukan alternatif yang optimal yaitu operator target program, selanjutnya setelah itu dilakukan tes pengujian akurasi sistem untuk membuktikan bahwa sistem pengambilan keputusan tersebut akurat dan juga dilakukan tes pengujian terhadap user apakah sistem tersebut telah sesuai telah sesuai dengan yang diharapkan. Hasil penelitian ini adalah dimana penentuan kecamatan yang dilakukan oleh PAM JAYA DKI menjadi lebih terkomputerisasi dan lebih objektif.

Kata Kunci : Sistem Penunjang Keputusan, *Cluster*, *Simple Additive Weighting (SAW)*, *Emerging Pattern*.

ABSTRACT

Emerging Pattern is one of the programs contained in MP3KI. Emerging Pattern intended to help accelerate poverty reduction throughout Indonesia. The central government and local governments jointly run the poverty program at district level poverty rates quite high. Indonesia has a total of 6994 districts and will continue to grow in line with the expansion in each region. With a total districts are pretty much the budget allocation for the program as well as a limited, the need for early selection to choose districts that will be the target of this Emerging Pattern program. Where each year the number will go up with the increase in the budget allocation. With the selection of this makes the central government, represented by PAM JAYA in determining the location of the target sub-district takes long enough. Therefore, to determine which sub-districts are entitled to the program, we need a system that can help to decision-making or can be called a decision support system. One method that can be used to make a decision support system is the integration of the Fuzzy C-Means method (FCM) clustering, Xie-Beni index and Methods Simple Additive Weighting (SAW). This method was chosen because it is considered able to select the best alternative from a number of alternatives, in this case the alternative is that the prospective sub-districts are eligible to receive a program based on certain criteria. Assessment is done by determining the grouping or pengclustering of poverty data exist, then calculate the index XB (Xie-Beni) on each - each cluster that has been formed so that the calculation is known cluster of which the most well that can be used as an alternative to making a decision, after the done perancangan process of each cluster that will determine the optimal alternative is the target sub-district program, then once that is done the test system accuracy testing to prove that the system is accurate decision making and also performed tests whether the user testing of the system is appropriate in accordance with the expected . From the results of this study is expected that the determination of the target sub-districts conducted by PAM JAYA be more objective and more efficient.

Key Word : Sistem Penunjang Keputusan, Cluster, Simple Additive Weighting (SAW), Emerging Pattern.

KATA PENGANTAR

Segala puji syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis ini yang berjudul ” MODEL DATA MINING PELANGGAN PAM JAYA DKI JAKARTA DENGAN TEKNIK *EMERGING PATTERN* (Studi Kasus PAM JAYA DKI Jakarta)”. Tujuan dari penulisan tesis ini adalah sebagai salah satu syarat untuk menyusun tesis pada Program Studi Magister Ilmu Komputer, Universitas Budi Luhur.

Rasa dan ucapan terima kasih penulis persembahkan kepada semua pihak yang telah membantu penulis dalam menyusun penelitian tesis ini :

1. Ibu Rusdah, M.Kom dan Bapak Spits Warnars Harco Leslie Hendric, Ph.D. Selaku dosen pembimbing yang telah membantu dan mengarahkan penulis.
2. Perusahaan Air Minum PAM JAYA DKI Jakarta yang telah memberikan ijin penulis untuk melakukan penelitian terhadap data konsumen dan informasi yang terkait dengan penelitian ini.
3. Hj. Neneng Badriah *ibunda tercinta dan keluarga* yang selalu memberikan semangat, dukungan dan doa bagi penulis.
4. Drg. Puti Octanesyaria Syahdan istri tercinta yang selalu memberikan doa kasih sayang, dukungan, semangat dan segalanya tanpa henti - hentinya kepada penulis.
5. Rekan - rekan mahasiswa MKOM Universitas Budi Luhur Kelas XK terima kasih atas kebersamaan, kerja keras dan dukungan semangatnya.
6. Bapak Ibu Dosen pengampu mata kuliah di Program Studi Magister Ilmu Komputer Universitas Budi Luhur yang telah dengan sabar memberikan ilmu pengetahuan, pencerahan dan bimbingan dalam belajar.
7. Semua pihak yang telah membantu penulis yang tidak dapat disebutkan namanya satu persatu oleh penulis.

Penulis menyadari bahwa masih banyak kekurangan dari penelitian tesis ini, oleh karena itu penulis mengharapkan kritik dan saran yang bersifat membangun demi kesempurnaan penelitian tesis dimasa mendatang. Semoga tesis ini dapat memberikan manfaat dari keterbatasannya. Amin.



Jakarta 25 Juni 2016

Penulis

Mohammad Ichlas

DAFTAR ISI

Halaman

LEMBAR JUDUL DEPAN	
LEMBAR JUDUL DALAM	
LEMBAR PERNYATAAN	
LEMBAR PENGESAHAN TESIS	
ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
 BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Masalah Penelitian	2
1.2.1 Identifikasi Masalah	2
1.2.2 Batasan Masalah	2
1.2.3 Rumusan Masalah	3
1.3 Tujuan dan Manfaat penelitian	3
1.5 Sistematika Penulisan	4
1.6 Daftar Pengertian	5
 BAB II LANDASAN TEORI DAN KERANGKA KONSEP	
2.1 Tinjauan Pustaka	7
2.1.1 Pengertian Data Mining	8
2.1.2 Pengelompokan Data Mining	15
2.1.3 Decision Tree	18
2.1.4 Emerging Pattern (EP)	18

2.1.5	Ekstraksi Rule dari Decision Tree	20
2.2	Tinjauan Studi	20
2.3	Metode Pemilihan Sampel	22
2.4	Tinjauan Obyek Penelitian	30
2.4.1	Sejarah PAM JAYA DKI Jakarta.....	30
2.4.2	Struktur Organisasi.....	32
2.5	Pola Pikir Pemecahan Masalah	32
2.6	Hipotesis	35
 BAB III METODOLOGI DAN RANCANGAN PENELITIAN		
3.1	Metode Penelitian	34
3.1.1	Data Primer.....	36
3.1.2	Data Sekunder	36
3.2	Metode Pengumpulan Data	37
3.2.1	Observasi	37
3.2.2	Wawancara	37
3.2.3	Studi Pustaka	37
3.3	Instrumen Penelitian	38
3.4	Rancangan dan Pengujian data Mining	39
3.5	Langkah-langkah Penelitian	40
3.6	Jadwal Penelitian	42
 BAB IV PEMBAHASAN HASIL PENELITIAN		
4.1	Pembahasan	43
4.1.1	Pemahaman Bisnis	43
4.1.2	Pemahaman Data	44
4.1.3	Persiapan Data	49
4.1.4	Pemodelan	52
4.1.5	Evaluasi	55
4.2	Implikasi Penelitian	57
4.2.1	Aspek Sistem	57

4.2.2	Aspek Manajerial	58
4.2.3	Aspek Penelitian Lanjutan	58
4.3	Rencana Implementasi.....	58

BAB V PENUTUP

5.1	Kesimpulan	59
5.2	Saran	60

DAFTAR PUSTAKA.....	61
----------------------------	-----------

LAMPIRAN - LAMPIRAN	-
----------------------------------	----------

RIWAYAT HIDUP	63
----------------------------	-----------



DAFTAR GAMBAR

Gambar	Halaman
II.1. Bidang Ilmu Data Mining	10
II.2 Proses dari Data Mining	12
II.3 Proses Data Mining Menurut CRISP-DM	13
II.4 Struktur PAM JAYA	32
II.5 Kerangka Pemikiran Penentuan Operator Sasaran	33
III.1. Langkah – Langkah Penelitian	41
IV.1. Cakupan Wilayah PAM DKI Jakarta	45
IV.2. Relasi Antar Tabel	46
IV.3. Tahap 1 (Pemilihan File SQL).....	54
IV.4. Tahap 2 (list field yang ada didalam table).....	54
IV.5. Tahap 3 (<i>Emergin Pattern</i> hasil dari tahap 2).....	55
IV.5. <i>Emerging Pattern</i> (itemset groupKode, KIIIA)	56

DAFTAR TABEL

Tabel	Halaman
I.1 Daftar Pengertian	5
II.1 Contoh Sederhana penerapan EP	19
II.2 Contoh Jumping sederhana pada EP	19
II.3 Daftar Penelitian	21
III.1 Jadwal Penelitian	43
IV.1 Tabel tbl_tarif.....	47
IV.2 Tabel tbl_biaya_tetap.....	47
IV.3 Tabel tbl_ukuran.....	48
IV.4 Tabel tbl_bill_group.....	48
IV.5 Tabel tbl_biaya_pemeliharaan.....	49
IV.6 Tabel tbl_grafik_mstplg.....	49
IV.7 Tabel tbl_grafik_pemakaian	50
IV.8 <i>Table Structure data set</i>	52
IV.9 <i>Table data group operator</i>	52
IV.10 Hasil <i>Emerging Pattern</i> 9 group klasifikasi.....	58
IV.11 Rencana Implementasi	59

DAFTAR PUSTAKA

- (Bor 2006) Borgelt, Cristian. 2006. An Implementation FP-Growth Algorithm.
Diambil November 2012, dari <http://www.borgelt.net/papers/fpgrowth.pdf>
(diakses pada 12 July 2015).
- (Cas 2007) Castro, Jose. 2007. C4.5 Decision Tree Creation and Pruning. Diambil
November 2012 dari http://mlrg.cecs.ucf.edu/MLRG_documents/c4.5.pdf
(diakses pada 12 July 2015).
- (Con 2004) Connolly, Thomas. Begg, Carolyn. 2004. Database System A
Practical Approach, Implementation, and Management Fourth Edition:
Addison Wesley. (diakses pada 12 July 2015).
- (Gar 2012) Garner, R Stephen. WEKA: The Waikato Environment for Knowledge
Analisis. Diambil November 2012 dari
<http://www.cs.waikato.ac.nz/~ml/publications/1995/Garner95-WEKA.pdf>
(diakses pada 12 July 2015).
- (Geh 1998) Gehrke, Johannes. Ramakrishnan, Raghu. Ganti, Venkatesh. 1998.
RainForest – A Framework for Fast Decision Tree Construction of Large
Datasets. Diambil November 2012 dari
<http://www.cs.cornell.edu/johannes/papers/1998/vldb1998-rainforest.pdf>
(diakses pada 12 July 2015).
- (Gor 2011) Gorunescu Florin. 2011. Data Mining: Concept, Models and
Techniques. Springer-Verleg Berlin Heidelberg
- (Han 2007) Han, Jiawei. Pei, Jian. 2007. Mining Frequent Pattern by Pattern-
Growth: Metodology and Implication. Diambil November 2012 dari
<http://www.acm.org/sigs/sigkdd/explorations/issues/2-2-2000-12/han.pdf> .
(diakses pada 12 July 2015).