

**OPTIMALISASI KINERJA HASIL KOMPARASI ALGORITMA C4.5 DAN
NAIVE BAYES MENGGUNAKAN PARTICLE SWARM OPTIMIZATION
UNTUK DETEKSI DINI PENENTUAN STATUS GIZI PADA BALITA :
STUDI KASUS DINAS KESEHATAN KABUPATEN ALOR**

TESIS



Oleh:

Jon Idrison Molina
1311601668

PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

JAKARTA
2015

**OPTIMALISASI KINERJA HASIL KOMPARASI ALGORITMA C4.5 DAN
NAIVE BAYES MENGGUNAKAN PARTICLE SWARM OPTIMIZATION
UNTUK DETEKSI DINI PENENTUAN STATUS GIZI PADA BALITA :
STUDI KASUS DINAS KESEHATAN KABUPATEN ALOR**

TESIS

Diajukan untuk memenuhi salah satu persyaratan
memperoleh gelar Magister Ilmu Komputer (MKOM)



Oleh:
Jon Idrison Molina
1311601668

PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

JAKARTA
2015



**PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama Mahasiswa : Jon Idrison Molina
Nomor Induk Mahasiswa : 1311601668
Program Studi : Magister Ilmu Komputer
Konsentrasi : Teknologi Sistem Informasi
Jenjang Studi : Strata Dua (S2)

Menyatakan bahwa Tesis yang berjudul:

Optimalisasi Kinerja Hasil Komparasi Algoritma C4.5 dan Naïve Bayes Menggunakan Particle Swarm Optimization Untuk Deteksi Dini Penentuan Status Gizi Pada Balita : Studi Kasus Dinas Kesehatan Kabupaten Alor

1. Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan 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 aturan yang berlaku apabila di kemudian hari pernyataan ini tidak benar.

Jakarta, 12 Agustus 2015



(Jon Idrison Molina)



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama Mahasiswa : Jon Idrison Molina
Nomor Induk Mahasiswa : 1311601668
Konsentrasi : Teknologi Sistem Informasi
Jenjang Studi : Strata Dua (S2)
Judul Tesis : Optimalisasi Kinerja Hasil Komparasi Algoritma C4.5 dan Naïve Bayes Menggunakan Particle Swarm Optimization Untuk Deteksi Dini Penentuan Status Gizi Pada Balita : Studi Kasus Dinas Kesehatan Kabupaten Alor

Jakarta, 12 Agustus 2015

Tim Penguji:

Tanda Tangan:

Ketua,
(Dr. Ir. Nazori AZ, M.T)


.....

Anggota,
(Ir. Wendi Usino, MM., M.Sc., Ph.D)


.....

Pembimbing,
(Hari Soetanto, S.Kom, M.Sc)


.....

Ketua Program Studi


(Dr. Ir. Nazori AZ, M.T)

ABSTRAK

Salah satu upaya cukup penting terhadap peningkatan kualitas sumber daya manusia adalah upaya peningkatan status gizi. Status gizi merupakan salah satu faktor yang menentukan kualitas hidup dan produktifitas kerja. Peningkatan status gizi sangat diperlukan apalagi pada balita yang masih berada dalam masa *golden age*. Penilaian status gizi bisa menjadi salah satu metode untuk mencegah terjadinya permasalahan gizi. Status gizi balita akan diketahui dengan adanya penilaian status gizi pada balita. Penilaian status gizi sangat berperan dalam pemantauan gizi anak. Apabila penilaian status gizi dilakukan secara tepat dan akurat maka adanya tanda atau gejala gangguan pertumbuhan maupun perkembangan anak dapat diketahui secara dini sehingga bisa dilakukan penanggulangan dan pencegahan dalam upaya peningkatan status gizi pada balita. Serta terjadinya permasalahan gizi dapat diminimalisir. Untuk itu maka penentuan status gizi pada balita harus dilakukan secara cepat tepat dan akurat. Puskesmas sebagai pelaksana teknis Dinas Kesehatan, mempunyai tugas pokok melakukan pendataan dan penilaian status gizi balita serta menyerahkan hasil penilaian ke Dinas Kesehatan. Penentuan status gizi yang selama ini dilakukan adalah dengan cara melihat dan menghitung pada tabel standar antropometri penilaian status gizi balita sesuai SK. Menteri Kesehatan Nomor:1995/Menkes/SK/XII/2010. Namun Puskesmas sering lambat dalam menangani pasien dalam jumlah yang banyak dan juga keterbatasan sumber daya terutama di wilayah yang kekurangan tenaga medis. Sehingga ini berpengaruh juga terhadap hasil penilaian penentuan status gizi balita. Untuk itu Dinas Kesehatan perlu menyediakan sebuah metode yang optimal dalam penentuan status gizi balita untuk dipakai oleh Puskesmas dalam penentuan status gizi balita secara cepat, tepat dan akurat. Untuk itu peneliti telah melakukan penelitian dengan menggunakan teknik komparasi dalam data mining sehingga dalam penelitian ini algoritma yang digunakan dalam teknik komparasi adalah algoritma C4.5 dan *Naïve Bayes*, hasil dari komparasi yang akurasiannya lebih tinggi akan dioptimalkan dengan menggunakan algoritma PSO untuk deteksi dini status gizi pada balita. Hasil pengujian dengan mengukur kinerja kedua algoritma tersebut dengan menggunakan metode pengujian *Confusion Matrix*, *Precision* dan *Recall*, diketahui bahwa algoritma C4.5 memiliki *accuracy* tertinggi yaitu 72.13% diikuti oleh algoritma *Naïve Bayes* dengan *accuracy* 70.33%. Dari algoritma yang paling tinggi tingkat keakurasiannya yaitu algoritma C4.5 dioptimalkan menggunakan algoritma PSO hasilnya menaikkan nilai *accuracy* menjadi 80.04%

Kata Kunci: Data Mining, C4.5, *Naïve Bayes*, *Particle Swarm Optimization*, Komparasi, Optimasi.

ABSTRACT

One important effort towards improving the quality of human resources is to increase the nutritional status. Nutritional status is one of the factors that determine the quality of life and work productivity. Improved nutritional status is very necessary especially in infants who are still in the golden age. Assessment of nutritional status can be one of the methods to prevent nutritional problems. Nutrition status will be determined by the assessment of nutritional status in infants. Assessment of nutritional status plays an important role in the monitoring of child nutrition. If the assessment of nutritional status performed precisely and accurately the signs or symptoms of impaired growth and development of children can be known at an early stage so that the reduction and prevention can be done in improving the nutritional status of children under five. As well as the occurrence of nutritional problems can be minimized. To that end, the determination of the nutritional status of infants should be performed quickly and accurately. PHC as the technical Department of Health, has a fundamental duty to collect data and assessment of nutritional status and submit the results of the assessment to the Department of Health. Determination of nutritional status that has been done is to look at the table and calculate the standard anthropometric nutritional status assessment according to SK. Minister of Health No. 1995 / Menkes / SK / XII / 2010. However, PHC is often slow in dealing with patients in large numbers and also the limitations of resources, especially in areas understaffed. This thus also affect the results of the determination of nutritional status assessment. For the Department of Health needs to provide an optimal method in determining the nutritional status of children for use by Puskesmas in determining the nutritional status of children in a rapid, precise and accurate. To that researchers have conducted research using comparative techniques in data mining so that in this study the algorithms used in a comparative technique is the algorithm C4.5 and Naïve Bayes, the results of comparative higher accuracy will be optimized using PSO algorithm for early detection of nutritional status in toddlers. Results of testing by measuring the performance of the two algorithms are using testing methods Confusion Matrix, Precision and Recall, it is known that the C4.5 algorithm has the highest accuracy is 72.13%, followed by Naïve Bayes algorithm with 70.33% accuracy. From the highest level algorithm is an algorithm C4.5 Accuracy optimized using the results raise the PSO algorithm accuracy value becomes 80.04%.

Keywords: Data Mining, C4.5, Naive Bayes, Particle Swarm Optimization, Comparison, Optimization.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yesus Kristus, karena berkat dan hikmat-Nya sehingga penulis dapat menyelesaikan Tesis yang berjudul Optimalisasi Kinerja Hasil Komparasi Algoritma C4.5 dan Naïve Bayes Menggunakan Particle Swarm Optimization Untuk Deteksi Dini Penentuan Status Gizi Pada Balita : Studi Kasus Dinas Kesehatan Kabupaten Alor. Tujuan dari penulisan tesis ini adalah sebagai salah satu syarat untuk menyusun tesis pada Program Studi Magister Ilmu Komputer, Universitas Budi Luhur Jakarta.

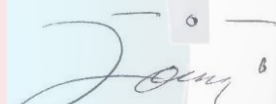
Rasa dan ucapan terima kasih penulis persembahkan kepada semua pihak yang telah membantu penulis dalam menyusun tesis ini:

1. Bapak Hari Soetanto, S.Kom, M.Sc, selaku dosen pembimbing tesis yang telah membimbing dan memotivasi penulis dalam mengerjakan tesis ini.
2. Bapak Kasih Hanggoro, MBA, Bapak Permenas L. Kolly, SE, Bapak Dr. Moedjiono, M.Sc yang sudah memberikan beasiswa untuk melanjutkan S2 di Universitas Budi Luhur.
3. Bapak Dr. Ir. Nazori AZ, M.T selaku Kaprodi MKom beserta Bapak dan 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.
4. Keluargaku, Bapak, Mama dan Kakak adik tercinta yang selalu mendoakan, memotivasi ku untuk menyelesaikan studi S2.
5. Kekasih ku tercinta *Gilberti Sailana, S.Pd.* yang selalu setia mendoakan, memotivasi ku dan tidak sabaran lagi menunggu ku meraih keberhasilan untuk mendapatkan gelar M.Kom.
6. Kakak sekaligus Teman-teman seperjuangan, *Pa Romly, Pa Mesak, Pa Letto, Pa Jun dan Pa Lasarus* yang selalu bersama dalam susah dan senang mengejar cita-cita kami di UBL, yang juga selalu mendoakan dan saling memotivasi untuk meraih kesuksesan bersama.
7. *Mas Yudo Devianto, Pa Kisworo, Mas Endah Wijaya* sebagai rekan sekaligus motivator yang selalu setia bersama-sama menyelesaikan tugas-tugas.

8. Rekan-rekan mahasiswa MKOM Universitas Budi Luhur kelas XB, XL Semester 1 dan 2 dan MKOM Semester 3 konsentrasi Teknologi Sistem Informasi, terima kasih atas kebersamaan, kerja keras dan dukungan semangatnya.
9. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu penulis selama studi hingga tugas akhir ini.

Penulis menyadari masih banyak kekurangan dari tugas akhir tesis ini, oleh karena itu penulis mengharapkan kritik dan saran yang bersifat membangun demi kesempurnaan penelitian tesis nantinya. Semoga tesis ini masih dapat memberikan manfaat dari sisi keterbatasannya.

Jakarta, 12 Agustus 2015



(Jon Idrison Molina)

HALAMAN PERSEMBAHAN

“Segala perkara dapat kutanggung di dalam Dia yang memberi kekuatan kepada ku” (Filipi 4:13)

*Kupersembahkan untuk kedua orang tua ku tercinta, keluarga besar ku
(Molina-Sesemamar-Susurmana)*



DAFTAR ISI

	Halaman
LEMBAR PERNYATAAN	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
KATA PENGANTAR	vi
HALAMAN PERSEMBAHAN.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xvi
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Masalah Penelitian	3
1.2.1 Identifikasi Masalah	3
1.2.2 Pembatasan Masalah.....	3
1.2.3 Rumusan Masalah.....	4
1.3. Tujuan dan Manfaat Penelitian	4
1.3.1 Tujuan Penelitian.....	4
1.3.2 Manfaat Penelitian.....	4
1.4. Tata Urut Penulisan.....	5
BAB II LANDASAN TEORI DAN KERANGKA KONSEP	7
2.1. Tinjauan Pustaka	7
2.1.1 Data Mining	7
2.1.2 Algoritma Klasifikasi Data Mining	11
2.1.3 Teknik - Teknik Data Mining.....	14
2.1.4 Algoritma Optimasi	19
2.1.5 Algoritma C4.5	20
2.1.6 Algoritma Naïve Bayes.....	26
2.1.7 Algoritma Particle Swarm Optimization (PSO)	30
2.1.8 Evaluasi dan Validasi Metode Klasifikasi Data Mining.....	32

2.1.9 Penentuan Status Gizi Pada Balita.....	34
2.1.10 RapidMiner.....	34
2.2. Tinjauan Studi.....	35
2.3. Tinjauan Obyek Penelitian.....	42
2.3.1 Profil Singkat Organisasi.....	42
2.3.2 Visi Misi Organisasi	42
2.3.3 Struktur Organisasi Dinas Kesehatan Kabupaten Alor	43
2.3.4 Sarana dan Prasarana SI Pendukung Penelitian	45
2.4. Kerangka Konsep	46
2.5. Hipotesis.....	47
BAB III METODOLOGI DAN RANCANGAN PENELITIAN	48
3.1. Jenis Penelitian.....	48
3.2. Populasi dan Sampel	48
3.2.1 Populasi	48
3.2.2 Sampel	48
3.3. Variabel Penelitian.....	49
3.4. Metode Pengumpulan Data.....	50
3.4.1 Studi Kepustakaan	50
3.4.2 Data Sekunder.....	50
3.5. Teknik Analisis Data.....	50
3.6. Eksperimen dan Pengujian Model	55
3.7. Evaluasi dan Validasi Hasil	57
3.8. Langkah - Langkah Penelitian	58
3.9. Jadwal Penelitian.....	60
BAB IV PEMBAHASAN HASIL PENELITIAN.....	62
4.1. Analisis Sistem.....	62
4.1.1 Analisis Proses Bisnis Sistem Berjalan	62
4.1.2 Analisis Kebutuhan Fungsional, Nonfungsional, dan Pengguna	63
4.2. Analisis Perilaku Sistem	67
4.2.1 Activity Diagram	67
4.2.2 Sequence Diagram.....	69

4.3. Perancangan Sistem	71
4.3.1 Deployment Diagram	71
4.3.2 Perancangan Infrastruktur Arsitektur	72
4.4. Konstruksi Sistem	74
4.4.1 Lingkungan Konstruksi	74
4.4.2 Konstruksi Database	74
4.5. Pengujian Sistem.....	78
4.5.1. Lingkungan Pengujian.....	78
4.5.2. Pengujian Validasi.....	79
4.5.3. Klasifikasi Algoritma <i>Naïve Bayes</i>	79
4.5.4. Klasifikasi Algoritma C4.5.....	82
4.6. Evaluasi dan Validasi.....	85
4.6.1. <i>Confusion Matrix</i> Algoritma <i>Naïve Bayes</i>	85
4.6.2. <i>Confusion Matrix</i> Algoritma C4.5.....	86
4.7. Analisa Hasil Komparasi.....	88
4.8. Optimasi Algoritma PSO dari Algoritma Terpilih.....	88
4.9. Penerapan Algoritma PSO	90
4.10. Implikasi Penelitian.....	91
4.10.1. Aspek Manajerial.....	91
4.10.2. Aspek Sistem	92
4.10.3. Rencana Implementasi.....	92
BAB V PENUTUP.....	94
5.1. Kesimpulan	94
5.2. Saran.....	94
DAFTAR PUSTAKA	96
LAMPIRAN-LAMPIRAN.....	99
DAFTAR RIWAYAT HIDUP.....	113

DAFTAR GAMBAR

Gambar	Halaman
II-1 Langkah-langkah dalam KDD.....	9
II-2 Contoh Metode <i>Clustering</i>	15
II-3 Contoh Metode <i>Association</i>	16
II-4 Contoh Kurva Metode <i>Forecasting</i>	18
II-5 Contoh penggunaan Metode <i>Sequence Analysis</i>	19
II-6 Pohon Keputusan Hasil Perhitungan Node 1	23
II-7 Pohon Keputusan Hasil Perhitungan Node 1.1	24
II-8 Pohon Keputusan Hasil Perhitungan Node 1.1.2	25
II-9 Bagan Struktur Organisasi Dinas Kesehatan Kab. Alor.....	43
II-10 Kerangka Pemikiran.....	47
III-1 Model yang diusulkan pada Rapid Miner 5.....	55
III-2 Metode cross validation pada Rapid Miner 5	56
III-3 Parameter untuk proses <i>decision tree</i>	57
III-4 Langkah-Langkah Penelitian	58
IV-1 Actor Model Sistem Deteksi Status Gizi Balita.....	66
IV-2 Use Case Diagram Sistem Deteksi Gizi Balita	66
IV-3 Activity Diagram Entry Master Nomor	68
IV-4 Sequence Diagram Entry Master Nomor.....	70
IV-5 Deployment Diagram Sistem Deteksi Status Gizi Balita	72
IV-6 Perancangan Infrastruktur	73
IV-7 Kontruksi Database Dengan Microsoft Access	74
IV-8 Tampilan Menu Login	75
IV-9 Tampilan Menu Utama Pengguna Admin	75
IV-10 Tampilan Entry Master Nomor	75
IV-11 Tampilan Entry Master User.....	76
IV-12 Tampilan Entry Master Gizi	76
IV-13 Tampilan Entry Master Tinggi Badan	76
IV-14 Tampilan Entry Master Umur.....	77

IV-15 Tampilan Entry Data Balita	77
IV-16 Tampilan Laporan Status Gizi Balita.....	78
IV-17 <i>Decision Tree</i> klasifikasi Status Gizi Balita dengan C4.5	85
IV-18 <i>Text View Model Confusion Matrix</i> untuk algoritma Naïve Bayes	86
IV-19 <i>Text View Model Confusion Matrix</i> untuk algoritma C4.5.....	87
IV-20 <i>Text View Model Confusion Matrix</i> algoritma C4.5 yang dioptimasi menggunakan PSO	89



DAFTAR TABEL

Tabel	Halaman
II-1 Contoh Metode <i>Classification</i>	14
II-2 Keputusan Bermain Tennis	21
II-3 Perhitungan Node 1	22
II-4 Perhitungan Node 1.1	23
II-5 Perhitungan Node 1.1.2	25
II-6 Data Pelatihan.....	28
II-7 Model <i>Confusion Matrix</i>	32
II-8 Penilaian status gizi balita menurut standar baku antropometri.....	34
II-9 Tinjauan Studi	40
III-1 Atribut dan nilai	49
III-2 Data Gizi Balita.....	51
III-3 Status Gizi Balita yang sudah melalui <i>processing</i> dan <i>cleansing</i> serta <i>ransformation</i> data.....	52
III-4 Jadwal Penelitian.....	60
IV-1 Tingkatan Pengguna dan Hak Akses	65
IV-2 Deskripsi <i>Use Case Diagram</i> Model Sistem Deteksi Gizi Balita	67
IV-3 Software Yang Digunakan.....	73
IV-4 Perhitungan Probabilitas <i>Prior</i>	79
IV-5 Perhitungan untuk Menentukan Klasifikasi Data <i>testing</i> X.....	81
IV-6 Hasil perhitungan <i>entropy</i> dan <i>gain</i> untuk menentukan simpul akar	83
IV-7 Model Tabel <i>Confusion Matrix</i> untuk algoritma <i>Naïve Bayes</i>	85
IV-8 Model Tabel <i>Confusion Matrix</i> untuk algoritma C4.5	87
IV-9 Komparasi Nilai Accuracy, Precision, dan Recall algoritma <i>Naïve Bayes</i> dan C4.5.....	88
IV-10 Model Tabel <i>Confusion Matrix</i> algoritma C4.5 yang dioptimasi menggunakan PSO	88
IV-11 Komparasi Nilai Accuracy, Precision, dan Recall algoritma C4.5 dan C4.5 + PSO.....	89

IV-12 Data Baru untuk Penerapan Algoritma Terpilih	90
IV-13 Confusion Matrix data baru dengan algoritma C4.5.....	91
IV-14 Spesifikasi <i>Hardware</i> dan <i>Software</i>	92
IV-15 Rencana Implementasi Sistem	92



DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Training.....	99
2. Data Baru untuk Penerapan Algoritma Terpilih.....	101
3. Activity Diagram Entry Master User.....	102
4. Activity Diagram Entry Master Gizi	103
5. Activity Diagram Entry Master Tinggi Badan	103
6. Activity Diagram Entry Master Berat Badan	104
7. Activity Diagram Entry Master Umur.....	104
8. Activity Diagram Entry Data Balita	105
9. Activity Diagram Pengolahan Laporan	105
10. Sequence Diagram Entry Master User	106
11. Sequence Diagram Entry Master Gizi	107
12. Sequence Diagram Entry Master Tinggi Badan	108
13. Sequence Diagram Entry Master Berat Badan	109
14. Sequence Diagram Entry Master Umur.....	110
15. Sequence Diagram Entry Data Balita.....	111
16. Sequence Diagram Pengolahan Laporan.....	112

DAFTAR PUSTAKA

- [Chamidah A. N. 2009] Chamidah, A N. 2009, Deteksi Dini Gangguan Pertumbuhan dan Perkembangan Anak. *Jurnal Pendidikan Khusus* Vol 1.
- [Octaviani dkk 2008] Octaviani, Ulfa., N. Juniarti., dan A. Mardiyah, 2008, Hubungan Keaktifan Keluarga Dalam Kegiatan Posyandu Dengan Status Gizi Balita Di Desa Rancaekek Kulon Kecamatan Rancaekek. Universitas Padjajaran. Bandung.
- [Supariasa dkk 2001] Supariasa I. D. N., B. Bakri., dan I. Fajar, 2001, *Penilaian Status Gizi*, EGC. Jakarta.
- [Turban dkk 2005] Turban, E., Aronson, J.E., & Liang, T., 2005, *Decision Support Sistem and Intelligent Sistem*.
- [Larose 2005] Larose, Daniel, T., 2005, *Discovering Knowledge in Data: An Introduction to Data Mining*. New Jersey : Jhon Willey & Son Inc.
- [Mohammed & Wagner 2014] Mohammed J, Zaki., & Wagner, Meira, Jr., 2014, *Data Mining And Analysis : Fundamental Concepts and Algorithms*. Rensselaer Polytechnic Institute, Troy, New York & Universidade Federal de Minas Gerais, Brazil
- [Sani & Dedy 2010] Sani, S., Dedy, S., Pengantar Data Mining: Menggali Pengetahuan dari Bongkahan Data, 2010.
- [Usama dkk 1996] Usama, Fayyad., Gregory, P.S., & Padhraic, Smyth., 1996, *From Data Mining to Knowledge Discovery in Databases*. American Association for Artificial Intelligence (AAAI).
- [Brachman & Anand 1996] Brachman, R., and Anand, T., 1996, *The Process of Knowledge Discovery in Databases: A Human-Centered Approach*. In *Advances in Knowledge Discovery and Data Mining*. Uthurusamy. Menlo Park, Calif.: AAAI Press.
- [Quinlan 1993] Quinlan, J. R., 1993, "*C4.5: Programs for Machine Learning*", California: Morgan Kaufmann.
- [Wu & Kumar 2009] Wu., & Kumar, V., 2009, *The Top Ten Algorithms in Data Mining*. Boca Raton : Chapman & Hall.