



Sekolah Tinggi Teknik Surabaya
Jalan Ngagel Jaya Tengah 73 - 77, Surabaya 60284, Indonesia
Telp. (031) 5027920 Fax. (031) 5041509

F-049

PROPOSAL TESIS

Semester Gasal / Genap *) Tahun ajaran _____ / _____

Program Studi : S2 - TEKNOLOGI INFORMASI
Nama Mahasiswa : TAJUDIN NOOR
NRP Mahasiswa : 212 210 313
Judul Tesis :



EVALUASI UMPAN BALIK KINERJA DOSEN
JURUSAN ADMINISTRASI BISNIS POLITEKNIK NEGERI
BANJARMASIN DENGAN METODE SUPPORT VECTOR MACHINES

Jenis Tesis : ☐ Kontribusi Keilmuan ☒ Kontribusi Pemanfaatan
Pembimbing Utama : HARTARTO JUNAEDI
Co. Pembimbing : _____
Jumlah SKS SUDAH LULUS : _____ SKS IPK : _____

Mengetahui,
Pembimbing Utama : _____ Co. Pembimbing : _____
Surabaya, _____
Pemohon : _____
(HARTARTO JUNAEDI, S.Kom, M.Kom) (TAJUDIN NOOR)

Catatan tambahan proposal (DISETUJUI / DIPERBAIKI / DITOLAK *)

KEPUTUSAN USULAN : DISETUJUI / DIPERBAIKI / DITOLAK *)

Pembimbing Utama : _____ Co. Pembimbing : _____

Surabaya, 13 - 11 - 2013

Mengetahui,

Pembantu Ketua I

Ketua Program Studi

13/13
11

(GUNAWAN)

(an. Jon, Erickson)

*) Coret yang tidak perlu

BAA - (02-02-2011) - Proposal Tesis v3 F-049.doc



SEKOLAH TINGGI TEKNIK SURABAYA

TERAKREDITASI Badan Akreditasi Nasional Perguruan Tinggi
Jl. Ngagel Jaya Tengah 73-77 Surabaya 60284, P.O.Box 1398, Indonesia
Telp. (031) 5027920 (Hunting); Fax. (031) 5041509
Web: <http://www.stts.edu>

CHECKLIST PROPOSAL TESIS S2-TEKNOLOGI INFORMASI

Ketentuan umum yang harus dipenuhi:

- ☒ Hipotesa tesis bukan merupakan fakta melainkan hasil penalaran yang jelas dan terukur, sedapat mungkin dapat diukur secara kuantitatif.
- ☒ Tinjauan Pustaka minimal 3 (tiga) paper yang menunjang dan/atau mendasari penelitian (diusahakan paper internasional), masing-masing tinjauan pustaka dibahas sepanjang 1 (satu) halaman di proposal.
- ☒ Block diagram proses beserta spesifikasi input-output sudah dijelaskan.
- ☒ Ruang lingkup minimal 2 (dua) halaman dan berisi batasan-batasan penelitian yang ditulis dengan jelas dan bukan merupakan teori-teori dasar.
- ☒ Target keberhasilan tesis sudah dijelaskan.
- ☒ Daftar pustaka minimal 5 paper, termasuk 3 judul paper di tinjauan pustaka.

Untuk kontribusi keilmuan:

- ☐ Dataset yang akan digunakan sudah dijelaskan spesifikasi dan jumlah datanya.
- ☐ Peningkatan dan/atau perbaikan yang akan dilakukan sudah dijelaskan.

Untuk kontribusi pemanfaatan:

- ☒ Spesifikasi Institusi tempat penelitian ini dimanfaatkan harus ditulis dengan jelas. (jumlah karyawan, jumlah cabang, jumlah komputer, jumlah jaringan, dan lain-lain tergantung tujuan tesis).
- ☒ Jumlah dan spesifikasi data yang akan digunakan dalam ujicoba keberhasilan pemanfaatan tesis sudah jelas.

*) Dosen pembimbing wajib memberikan check (✓) untuk tiap point yang telah dipenuhi.

Surabaya, _____

TADJUDIN NOOR

NRP: 212 210 313

Dosen Pembimbing

PROPOSAL TESIS

**EVALUASI UMPAN BALIK KINERJA DOSEN
JURUSAN ADMINISTRASI BISNIS
POLITEKNIK NEGERI BANJARMASIN
DENGAN METODE SUPPORT VECTOR MACHINES**



TAJUDIN NOOR
212210313

PROGRAM PASCASARJANA
MAGISTER TEKNOLOGI INFORMASI
SEKOLAH TINGGI TEKNIK SURABAYA
JUNI 2013

A-1

Judul : Evaluasi Umpan Balik Kinerja Dosen
Jurusan Administrasi Bisnis Politeknik
Negeri Banjarmasin Dengan Metode Support
Vector Machines
Nama : Tajudin Noor
NRP : 212210313
Program Studi : Magister Teknologi Informasi
Dosen pembimbing : Hartarto Junaedi, S.Kom, M.Kom.
Co. Pembimbing : -
Jenis Tesis : Kontribusi Pemanfaatan

I. Latar Belakang

Tercapainya tujuan proses belajar mengajar dalam suatu perguruan tinggi tidak terlepas dari peranan dosen dan mahasiswa. Keaktifan para dosen dalam memberikan perkuliahan dan keaktifan mahasiswa dalam mengikuti proses belajar mengajar menjadi kunci utama suksesnya proses belajar mengajar. Suksesnya proses belajar mengajar, bagi mahasiswa, dapat dilihat dengan alat ukur berupa nilai akhir yang diperoleh. Biasanya, seorang mahasiswa dikatakan memiliki nilai baik dalam suatu matakuliah, apabila mahasiswa tersebut mendapatkan nilai lebih dari atau sama dengan 'B'. Demikian pula, seorang dosen dikatakan sukses dalam proses belajar mengajar, apabila nilai kinerja yang diperolehnya juga baik¹.

Kualitas dosen memegang peranan penting di suatu perguruan tinggi yang ingin mencapai tujuan proses belajar mengajar yaitu menghasilkan lulusan (*output*) yang berkualitas. Dosen dinilai berkualitas apabila memiliki nilai kinerja yang baik, yang ditinjau dari beberapa aspek. Untuk itu sangat perlu dilakukan sebuah analisis untuk mengetahui nilai kinerja dari seorang dosen khususnya di Program Studi Manajemen Informatika Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin yang nantinya akan dijadikan sebagai masukan bagi bagian akademik Politeknik Negeri Banjarmasin.

Sebelumnya Program Studi Manajemen Informatika Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin telah menerapkan evaluasi terhadap kinerja dosen dengan menggunakan alat ukur yang dikenal dengan nama Evaluasi Umpan Balik (EUB). Salah satu variabel yang digunakan untuk menghitung nilai kinerja dosen adalah penilaian dari mahasiswa yang mengambil matakuliah yang diampu oleh dosen yang bersangkutan. Penilaian oleh mahasiswa dilakukan melalui

¹ Sundani, 2008.

pengisian kuesioner. Kendati hal ini sudah dilakukan, namun hasil yang dicapai kurang maksimal. Mengingat hasil dari evaluasi umpan balik yang selama ini sudah dilakukan belum menghasilkan nilai akhir dari kinerja dosen.

Pada semester ganjil 2011/2012, ada 14 pertanyaan yang diberikan dalam kuesioner, yaitu: kemampuan pengajar dalam merangsang dan menumbuhkan minat terhadap matakuliah, kepedulian pengajar yang ditujukan pada minat terhadap matakuliah, kemampuan berkomunikasi dengan cara yang bisa dimengerti, kesempatan untuk mengajukan pertanyaan tentang materi kuliah, pengorganisasian materi perkuliahan dari matakuliah, cara penyampaian materi perkuliahan, rangsangan untuk berpikir kritis dan analitis, relevansi tugas yang diberikan terhadap matakuliah, kesediaan pengajar untuk konsultasi dengan mahasiswa, pengembalian kertas ujian dan tugas. Setiap pertanyaan memiliki nilai mulai dari 1 (Kurang Sekali), 2 (Kurang), 3 (Cukup), 4 (Baik), 5 (Baik Sekali). Penilaian dari mahasiswa ini bersifat kualitatif. Selain penilaian mahasiswa, penilaian kinerja dosen juga dipengaruhi oleh kehadiran dosen yang bersifat kuantitatif. Pada penelitian ini dikembangkan sistem kuesioner penilaian evaluasi umpan balik kinerja dosen dengan menggunakan *text mining* sebagai otomatisasi dalam penggalian informasi dari data kuesioner mahasiswa yang berupa teks. Otomatisasi kuesioner dibuat dengan menggunakan *free text*, tanpa adanya pilihan jawaban “ya”, atau “tidak”, “cukup” atau “kurang” ataupun yang lainnya. *Free text* digunakan karena beberapa pertimbangan, yaitu (1) menghindari terjadinya kesalahan persepsi dalam kuesioner, (2) lebih leluasa dalam mengungkapkan pendapat oleh mahasiswa dan (3) menghindari manipulasi keputusan yang dibuat.² *Text mining*, mengacu pada proses mengambil informasi berkualitas tinggi dari teks. Informasi berkualitas tinggi biasanya diperoleh melalui peramalan pola dan kecenderungan melalui sarana seperti pembelajaran pola statistik. Proses *text mining* yang khas meliputi kategorisasi teks, *text clustering*, ekstraksi konsep/entitas, produksi taksonomi granular, *sentiment analysis*, penyimpulan dokumen, dan pemodelan relasi entitas (yaitu, pembelajaran hubungan antara entitas bernama) (Wikipedia, 2011).

Teknik klasifikasi yang telah memperoleh perhatian serius adalah *Metode Support Vector Machines* (SVM). Teknik ini berakar pada teori pembelajaran statistik dan telah menunjukkan hasil empiris yang menjanjikan dalam berbagai aplikasi praktis dari pengenalan digit tulisan tangan sampai kategorisasi teks. SVM juga bekerja sangat baik pada data dengan banyak dimensi dan menghindari kesulitan dari permasalahan dimensionalitas (Pang-Ning Tan, dkk, 2006).

Metode tersebut banyak digunakan dalam kategorisasi teks. Pada hasil eksperimen (Wulandini, F. & Nugroho, A. N. 2009) untuk kategorisasi teks berbahasa

² Mattison R, 1999 dan Sullivan D, 2001

Indonesia didapatkan bahwa SVM menunjukkan performansi yang sedikit lebih baik dengan akurasi 92,5% dibandingkan metode NBC dengan akurasi 90% padahal metode NBC adalah metode yang jauh lebih konvensional dan lebih sederhana. Beberapa metode lain yang digunakan pula untuk proses teks mining adalah C45, K-Nearest Neighbor, K-Means dan algoritma genetika. Dari hasil penelitian (Wulandini, F. & Nugroho, A. N. 2009) untuk kategorisasi teks berbahasa Indonesia didapatkan hasil akurasi 29,17% untuk *K-Nearest Neighbor* dan 7,5% untuk metode C45. Oleh karena itu pada penelitian ini dipilih metode NBC untuk menyelesaikan masalah *sentiment analysis*.

Sentiment Analysis atau *opinion mining* adalah studi komputasional dari opini-opini orang, *appraisal* dan emosi melalui entitas, *event* dan atribut yang dimiliki (Biu, L. 2010). Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur/tingkat aspek - apakah pendapat yang dikemukakan dalam dokumen, kalimat atau fitur entitas / spek bersifat positif, negatif atau netral (Dehaff, M., 2010).

II. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, didapat suatu perumusan masalah sebagai berikut :

Bagaimana menerapkan *text mining* dengan *Metode Support Vector Machine* (SVM) pada evaluasi umpan balik kinerja dosen pada Program Studi Manajemen Informatika Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin sehingga evaluasi tersebut mampu menghasilkan opini setiap penilaian kinerja dosen.

III. Tujuan

Tujuan yang ingin dicapai pada penelitian tesis ini adalah:

1. Menggali pengetahuan (discovering knowledge) tentang pola (pattern) proporsi pelaksanaan proses belajar mengajar oleh dosen sehingga diketahui feedback dari mahasiswa untuk produktivitas, komposisi serta keseimbangan dalam pelaksanaan pembelajaran tersebut.
2. Memberi nilai tambah dengan menambah unsur *text mining* terhadap program Beban Kerja Dosen dan Evaluasi Pelaksanaan Tridharma Perguruan Tinggi.
3. Pengembangan sebuah software yang dapat mengklasifikasi opini positif, sangat positif, negatif, sangat negatif dari kuesioner umpan balik mahasiswa menggunakan *Metode Support Vector Machine* (SVM) serta menganalisis performansi metode tersebut.

IV. Hipotesa

Hipotesa yang diambil pada penelitian ini adalah sebagai berikut :
Dengan menggunakan *Metode Support Vector Machines* pada pengklasifikasian data kuesioner umpan balik kinerja dosen Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin akan didapatkan tingkat akurasi minimal sebesar 80 %.

V. Tinjauan Pustaka

Tinjauan Pustaka yang digunakan adalah :

1. Opini Mining³

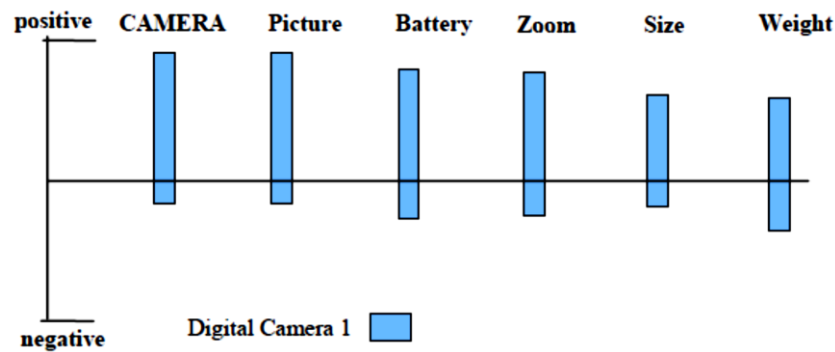
Penelitian ini bertujuan untuk mengetahui opini masyarakat terhadap produk kamera digital, dengan mengklasifikasikan menjadi 2 data output klasifikasi yaitu positif dan negatif. Ada beberapa data yang berupa data yang menjadi pertimbangan opini antara lain : camera, kualitas gambar dan ukuran, dari sampel yang diberikan oleh Bing Liu sebagai berikut :

Tabel 1. Hasil klasifikasi produk kamera

No	Kategori	Positif	Negatif
1	Camera	125	7
2	Kualitas Gambar	123	6
3	Ukuran	82	10
	Jumlah	330	23

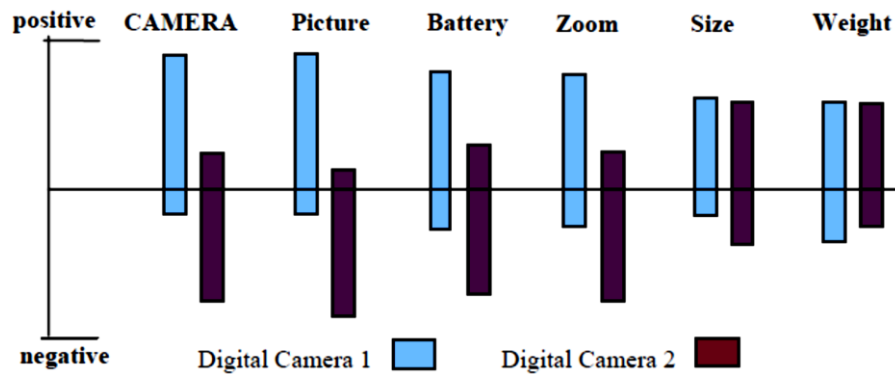
Untuk selain hasil dari segi kriteria diatas pada tabel diatas maka dapat ditambahkan beberapa katagori sebagai pengklasifikasiannya misalnya untuk ketahanan battery, kemampuan zoom dan lain sebagainya. Maka pada paper Bing Liu dapat di rpresentasikan dalam bentuk grafik pada gambar 1 dan gambar 2 di bawah ini.

³ Bing Liu, *Opinion Mining*. Department of Computer Science University of Illinois at Chicago
851 S. Morgan Street Chicago, IL 60607-0753



Gambar 1. hasil grafik klasifikasi terhadap 1 kamera

Pada gambar 1 grafik digambarkan dengan arah vertical dimana semakin banyak chart ke arah positif itu menandakan persentase tiap masing-masing kategori yang diberikan hampir lebih ke arah positif.



Gambar 2. hasil grafik klasifikasi terhadap 2 kamera

Dari gambar 2 ini menunjukkan ketika ditambahkan satu buah objek kamera lain sebagai pembandingan. Untuk *chart* warna biru menunjukkan sentiment terhadap kamera pertama dan *chart* warna ungu menunjukkan sentiment terhadap kamera ke dua. Setiap kamera memiliki kategori yang sama sebagai klasifikasinya misalnya, dilihat dari segi atau kelompok hasil gambar, ketahanan *battery*, kemampuan *zoom*, dan ukuran kamera serta berat dari kamera tersebut.

Penelitian Bing Liu menghasilkan bahwa produk kamera digital adalah positif atau bagus dalam penilaian konsumen. Klasifikasi opini yang dibuat Bing Liu termasuk dalam kategori *Fined-grained sentiment analysis*, karena obyek yang dianalisis adalah kalimat.

2. Text Classification Using Support Vector Machine for Webmining Based Spatio Temporal Analysis of the Spread of Tropical Diseases ⁴

Pada penelitian bertujuan menggali informasi jenis penyakit pada alam tropis seperti demam berdarah, malaria serta flu burung. Informasi digali dari website dengan menggunakan data mining untuk mengekstrak informasi. Dengan web mining informasi penyakit di iklim tropis dikumpulkan. Penelitian ini dikembangkan dengan klasifikasi teks menggunakan Metode Support Vector Machines, klasifikasi yang tepat akan mampu mengekstraksi informasi dengan baik selanjutnya dapat menganalisis dan memvisualisasikan. *Metode Support Vector Machines* mempunyai kapabilitas untuk mengklasifikasi data teks hingga pada skala yang besar.

Penelitian ini mendapatkan hasil bahwa *Metode Support Vector Machines* menunjukkan hasil paling baik pada kategorisasi teks berbahasa Indonesia. Pada eksperimen yang dilakukan pada 3713 *features* dan 360 *instances*. Data tersebut dibagi menjadi 120 *instances* sebagai data uji dan 240 *instances* sebagai data latih. Hasilnya seperti ditunjukkan oleh tabel 1 sebagai berikut :

Tabel 2.

Hasil Eksperimen Fatimah Wulandini dan Anto Satriyo Nugroho

Metode	Akurasi
SVM	92,5 %
K- Nearest Neighbor	49,17 %
Naïve Bayes Classifier	90 %
C45	77,5 %

Tabel 1 tersebut menunjukkan performansi dari SVM di bandingkan dengan Metode *K- Nearest Neighbor*, *Naïve Bayes Classifier* dan *C45* yang akurasiya diatas dari metode-metode tersebut.

⁴ Fatimah Wulandini dan Anto Satriyo Nugroho, *Text Classification Using Support Vector Machine for Webmining Based Spation Temporal Analysis of the Spread of Tropical Diseases*, 2009

3. Improving Multiclass Text Classification with the Support Vector Machine⁵

Penelitian ini membandingkan *Metode Naive Bayes* dan *Metode Support Vector Machines* pada *multiclass* klasifikasi teks. Menggunakan pendekatan kombinasi klasifikasi *biner*, hasilnya secara substansial *SVM* mengungguli *Naive Bayes*. Dengan menggunakan *full multiclass* pada dua set data teks terkenal, sampai pada margin *error* terendah pada kedua set data tersebut. Dikembangkan sebuah indikator baru dari kinerja biner untuk menunjukkan pada *lower multiclass SVM* menghasilkan peningkatan pada kinerja biner.

Menunjukkan hasil bahwa *SVM* menghasilkan performansi yang lebih baik dalam menyelesaikan klasifikasi teks multi kelas dibandingkan *Metode NBC*. Hal ini sesuai dengan pernyataan *Fabrice Colas & Pavel Brazdil* bahwa *SVM* unggul dalam klasifikasi *multiclass*.

Dibawah ini adalah tabel hasil perbandingan *multiclass* pada 20 dataset koran dan sektor industri. Pada baris atas menunjukkan jumlah dokumen (data training yang digunakan), sedangkan pada baris kedua menunjukkan metode klasifikasi.

Tabel 3.
Perbandingan Metode SVM dan NBC

20 Newsgroups	800		250		100		30	
	SVM	NB	SVM	NB	SVM	NB	SVM	NB
OVA	0.131	0.146	0.167	0.199	0.214	0.277	0.311	0.445
Dense 15	0.142	0.176	0.193	0.222	0.251	0.282	0.366	0.431
BCH 15	0.145	0.169	0.196	0.225	0.262	0.311	0.415	0.520
Dense 31	0.135	0.168	0.180	0.214	0.233	0.276	0.348	0.428
BCH 31	0.131	0.153	0.173	0.198	0.224	0.259	0.333	0.438
Dense 63	0.129	0.154	0.171	0.198	0.222	0.256	0.326	0.407
BCH 63	0.125	0.145	0.164	0.188	0.213	0.245	0.312	0.390

Industry Sector	52		20		10		3	
	SVM	NB	SVM	NB	SVM	NB	SVM	NB
OVA	0.072	0.357	0.176	0.568	0.341	0.725	0.650	0.885
Dense 15	0.119	0.191	0.283	0.363	0.461	0.542	0.738	0.805
BCH 15	0.106	0.182	0.261	0.352	0.438	0.518	0.717	0.771
Dense 31	0.083	0.145	0.216	0.301	0.394	0.482	0.701	0.769
BCH 31	0.076	0.140	0.198	0.292	0.371	0.462	0.676	0.743
Dense 63	0.072	0.135	0.189	0.279	0.363	0.453	0.674	0.745
BCH 63	0.067	0.128	0.176	0.272	0.343	0.443	0.653	0.734

⁵ Jason D. M. Rennie & Ryan Rifkin, *Improving Multiclass Text Classification with the Support Vector Machine*

4. Text Mining⁶

Text mining, mengacu pada proses mengambil informasi berkualitas tinggi dari teks. Informasi berkualitas tinggi biasanya diperoleh melalui peramalan pola dan kecenderungan melalui sarana seperti pembelajaran pola statistik. *Text mining* biasanya melibatkan proses penataan teks input (biasanya parsing, bersama dengan penambahan beberapa fitur linguistik turunan dan penghilangan beberapa diantaranya, dan penyisipan *subsequent* ke dalam database), menentukan pola dalam data terstruktur, dan akhirnya mengevaluasi dan menginterpretasi output. 'Berkualitas tinggi' di bidang *text mining* biasanya mengacu ke beberapa kombinasi relevansi, kebaruan, dan *interestingness*. Proses *text mining* yang khas meliputi kategorisasi teks, *text clustering*, ekstraksi konsep/entitas, produksi taksonomi granular, *sentiment analysis*, penyimpulan dokumen, dan pemodelan relasi entitas (yaitu, pembelajaran hubungan antara entitas bernama). Klasifikasi / kategorisasi dokumen adalah masalah dalam ilmu informasi. Tugas kita adalah untuk menetapkan dokumen elektronik masuk dalam satu atau lebih kategori, berdasarkan isinya. Tugas klasifikasi dokumen dapat dibagi menjadi dua macam yaitu klasifikasi dokumen terawasi di mana beberapa mekanisme eksternal (seperti feedback manusia) memberikan informasi mengenai klasifikasi yang tepat untuk dokumen, dan klasifikasi dokumen tak terawasi, dimana klasifikasi harus dilakukan sepenuhnya tanpa merujuk ke informasi eksternal. Ada juga klasifikasi dokumen semi-diawasi, dimana bagian dari dokumen diberi label oleh mekanisme eksternal.

Pendekatan manual *text mining* secara intensif dalam laboratorium pertama muncul pada pertengahan 1980-an, namun kemajuan teknologi telah memungkinkan ranah tersebut untuk berkembang selama dekade terakhir. *Text mining* adalah bidang interdisipliner yang mengacu pada pencarian informasi, pertambangan data, pembelajaran mesin, statistik, dan komputasi linguistik. Dikarenakan kebanyakan informasi (perkiraan umum mengatakan lebih dari 80%) saat ini disimpan sebagai teks, *text mining* diyakini memiliki potensi nilai komersial tinggi (Clara Bridge, 2011).

Saat ini , *text mining* telah mendapat perhatian dalam berbagai bidang :

a. Aplikasi keamanan.

Banyak paket perangkat lunak *text mining* dipasarkan terhadap aplikasi keamanan, khususnya analisis *plain text* seperti berita internet. Hal ini juga mencakup studi enkripsi teks.

b. Aplikasi biomedis.

⁶ Wikipedia, 2011

Berbagai aplikasi *text mining* dalam literatur biomedis telah disusun. Salah satu contohnya adalah PubGene yang mengkombinasikan *text mining* biomedis dengan visualisasi jaringan sebagai sebuah layanan Internet. Contoh lain *text mining* adalah GoPubMed.org. Kesamaan semantik juga telah digunakan oleh sistem *text mining*, yaitu, GOAnnotator.

c. Perangkat Lunak dan Aplikasi

Departemen riset dan pengembangan perusahaan besar, termasuk IBM dan Microsoft, sedang meneliti teknik *text mining* dan mengembangkan program untuk lebih mengotomatisasi proses pertambangan dan analisis. Perangkat lunak *text mining* juga sedang diteliti oleh perusahaan yang berbeda yang bekerja di bidang pencarian dan pengindeksan secara umum sebagai cara untuk meningkatkan performansinya.

d. Aplikasi Media Online

Text mining sedang digunakan oleh perusahaan media besar, seperti perusahaan Tribune, untuk menghilangkan ambiguitas informasi dan untuk memberikan pembaca dengan pengalaman pencarian yang lebih baik, yang meningkatkan loyalitas pada site dan pendapatan. Selain itu, editor diuntungkan dengan mampu berbagi, mengasosiasikan dan properti paket berita, secara signifikan meningkatkan peluang untuk menguangkan konten.

e. Aplikasi Pemasaran

Text mining juga mulai digunakan dalam pemasaran, lebih spesifik dalam analisis manajemen hubungan pelanggan. Coussemont dan Van den Poel (2008) menerapkannya untuk meningkatkan model analisis prediksi untuk *churn* pelanggan (pengurangan pelanggan).

f. *Sentiment Analysis*

Sentiment Analysis mungkin melibatkan analisis dari review film untuk memperkirakan berapa baik review untuk sebuah film. Analisis semacam ini mungkin memerlukan kumpulan data berlabel atau label dari efektivitas kata-kata. Sebuah sumber daya untuk efektivitas kata-kata telah dibuat untuk WordNet.

g. Aplikasi Akademik

Masalah *text mining* penting bagi penerbit yang memiliki database besar untuk mendapatkan informasi yang memerlukan pengindeksan untuk pencarian. Hal ini terutama berlaku dalam ilmu sains, di mana informasi yang sangat spesifik sering terkandung dalam teks tertulis. Oleh karena itu, inisiatif telah diambil seperti *Nature's proposal* untuk *Open Text Mining Interface* (OTMI) dan *Health's common Journal Publishing* untuk *Document Type Definition* (DTD) yang akan memberikan isyarat semantik pada mesin untuk menjawab pertanyaan spesifik yang terkandung dalam teks tanpa menghilangkan *barrier* penerbit untuk akses publik.

Sebelumnya, website paling sering menggunakan pencarian berbasis teks, yang hanya menemukan dokumen yang berisi kata-kata atau frase spesifik yang ditentukan oleh pengguna. Sekarang, melalui penggunaan web semantik, *text mining* dapat menemukan konten berdasarkan makna dan konteks (daripada hanya dengan kata tertentu). *Text mining* juga digunakan dalam beberapa filter email spam sebagai cara untuk menentukan karakteristik pesan yang mungkin berupa iklan atau materi yang tidak diinginkan lainnya.

5. *Sentiment analysis*⁷

Sentiment analysis atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan *text mining*. Secara umum, bertujuan untuk menentukan *attitude* pembicara atau penulis berkenaan dengan topik tertentu. *Attitude* mungkin penilaian atau evaluasi mereka, pernyataan efektif mereka (pernyataan emosional penulis saat menulis) atau komunikasi emosional dimaksud (efek emosional penulis inginkan terhadap pembaca).

Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur / tingkat aspek – apakah pendapat yang dikemukakan dalam dokumen, kalimat atau fitur entitas / aspek bersifat positif, negatif atau netral (Dehaff, M., 2010). Lebih lanjut *sentiment analysis* dapat menyatakan emosional sedih, gembira, atau marah. Beberapa penelitian mengklasifikasikan polaritas dokumen pada skala multi-arah, yang dicoba oleh (Pang, B. & Lee, L. 2005) dan (Snyder B. & Barzilay R. 2007) antara lain : memperluas tugas dasar klasifikasi review film sebagai positif atau negatif terhadap memprediksi peringkat bintang baik skala 3 atau bintang 4, sementara (Snyder B. & Barzilay R. 2007) melakukan analisa mendalam tentang review restoran, memprediksi peringkat untuk berbagai aspek dari restoran yang diberikan, seperti makanan dan suasana (dalam skala bintang lima).

Sebuah metode yang berbeda untuk menentukan sentimen adalah penggunaan sistem skala dimana kata-kata umumnya terkait memiliki sentimen negatif, netral atau positif dengan mereka diberi nomor pada skala -5 sampai +5 (paling negatif hingga yang paling positif) dan ketika sepotong teks terstruktur dianalisis dengan pemrosesan bahasa alami, konsep selanjutnya dianalisis untuk memahami kata-kata ini dan bagaimana mereka berhubungan

⁷ Wikipedia, 2011

dengan konsep. Setiap konsep kemudian diberi skor berdasarkan bagaimana kata-kata sentimen berhubungan dengan konsep, dan skor yang terkait. Hal ini memungkinkan gerakan untuk pemahaman yang lebih canggih dari sentimen berdasarkan skala 11 titik. Penelitian dengan arah berbeda adalah identifikasi subjektivitas / objektivitas. Tugas ini biasanya didefinisikan sebagai menggolongkan suatu teks yang diberikan (biasanya kalimat) ke salah satu dari dua kelas: objektif atau subjektif (Pang, B. & Lee, L, 2008). Masalah ini kadang-kadang dapat lebih sulit daripada klasifikasi polaritas (Mihalcea, R. & dkk, 2007) subjektivitas kata-kata dan frase mungkin tergantung pada konteks dan dokumen objektif mungkin berisi kalimat subjektif (misalnya, sebuah artikel berita mengutip pendapat orang). Selain itu, seperti yang disebutkan oleh (Su, F. & Markert, K. 2008), hasilnya sangat tergantung pada definisi subjektivitas digunakan ketika memberikan keterangan pada teks. Namun, (Pang, B. & Lee, L. 2004) menunjukkan bahwa menghapus kalimat objektif dari sebuah dokumen sebelum mengelompokkan polaritasnya membantu meningkatkan kinerja.

Kita dapat melacak produk-produk, merek dan orang-orang misalnya dan menentukan apakah mereka dilihat positif atau negatif di web. Hal ini memungkinkan bisnis untuk melacak:

- a. Deteksi Flame (rants buruk)
- b. Persepsi produk baru.
- c. Persepsi Merek.
- d. Manajemen reputasi.

Hal ini juga memungkinkan individu untuk mendapatkan sebuah pandangan tentang sesuatu (review) pada skala global (Jenkins, M. C., 2011). Orang sering kali menyatakan lebih dari satu opini "*the movie was terrible, but DeNiro's performance was superb, as always*", sebuah sarkasme "*this is probably the best laptop Dell could come up with*", atau menggunakan negasi dan banyak elemen kompleks sehingga sulit untuk diparsing "*not that I'm saying this was a bad experience*". Ekspresi atau *sentiment* mengacu pada fokus topik tertentu, pernyataan pada satu topik mungkin akan berbeda makna dengan pernyataan yang sama pada *subject* yang berbeda. Sebagai contoh, adalah hal yang baik untuk mengatakan alur film tidak terprediksi, tapi adalah hal yang tidak baik jika 'tidak terprediksi' dinyatakan pada kemudi dari kendaraan. Bahkan pada produk tertentu, kata-kata yang sama dapat menggambarkan makna kebalikan, contoh adalah hal yang buruk untuk waktu *start-up* pada kamera digital jika dinyatakan "lama", namun jika "lama" dinyatakan pada usia baterai maka akan menjadi hal positif. Oleh karena itu pada beberapa penelitian, terutama pada review produk, pekerjaan didahului

dengan menentukan elemen dari sebuah produk yang sedang dibicarakan sebelum memulai proses *opinion mining* (Ian Barber, 2010).

Hal pertama dalam pemrosesan dokumen adalah memecah kumpulan karakter ke dalam kata atau token, sering disebut sebagai tokenisasi. Tokenisasi adalah hal yang kompleks untuk program komputer karena beberapa karakter dapat ditemukan sebagai *token delimiters*. Delimiter adalah karakter spasi, tab dan baris baru “*newline*”, sedangkan karakter () < > ! ? “ kadang kala dijadikan delimiter namun kadang kala bukan tergantung pada lingkungannya (Wulandini, F. & Nugroho, A. N. 2009).

VI. Ruang Lingkup

A. Data

1. Data Input

Data input yang digunakan adalah hasil kuesioner mahasiswa yang diinputkan kedalam sistem, sesuai dengan matakuliah yang diampu oleh dosen yang bersangkutan.

2. Desain Input Kuesioner Umpan Balik Evaluasi Kinerja Dosen Program Studi Manajemen Informatika Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin adalah sebagai berikut:

KUESIONER KINERJA AKADEMIK DOSEN	
JURUSAN/PRODI	:
MATA KULIAH (MK)	:
DOSEN PENGAMPU	:
*Jawab Pertanyaan dengan singkat dan jelas	
A. PROSES BELAJAR MENGAJAR	
1.	Apakah dosen telah merencanakan materi dan tujuan mata kuliah diberikan di awal perkuliahan ?
2.	Dosen datang tepat waktu & mengajar sesuai waktu yang terjadwal ?
3.	Dalam perkuliahan ada acara diskusi & tanya jawab ?
4.	Menurut anda, apakah ada manfaat soal latihan atau studi kasus dalam menambah pemahaman mata kuliah ini ?
5.	Sudah sesuai evaluasi (tugas dan UTS) dengan materi yang diajarkan ?
6.	Apakah ada pembahasan soal-soal, tugas dan UTS yang diberikan ?
7.	Bagaimana dosen dalam menjelaskan sistematika kuliah ?
8.	apakah dilakukan latihan soal terhadap setiap materi yang diberikan ?
9.	sesuai kah materi yang diberikan terhadap rencana di awal perkuliahan ?
B. KAPABILITAS (KOMPETENSI DOSEN)	
10.	Bagaimana kah kemampuan dosen dalam menjelaskan materi perkuliahan ?
11.	Menurut anda, bagaimana tingkat penguasaan materi, wawasan, dan implementasi mata kuliah ini untuk dosen yang bersangkutan ?
12.	Kemampuan dosen menjawab pertanyaan dari mahasiswa ?
13.	Semangat dosen dalam memberikan kuliah ?
14.	Kemampuan dosen dalam memberikan motivasi/membangkitkan minat belajar ?

Gambar 3. Desain Input Data Kuesioner

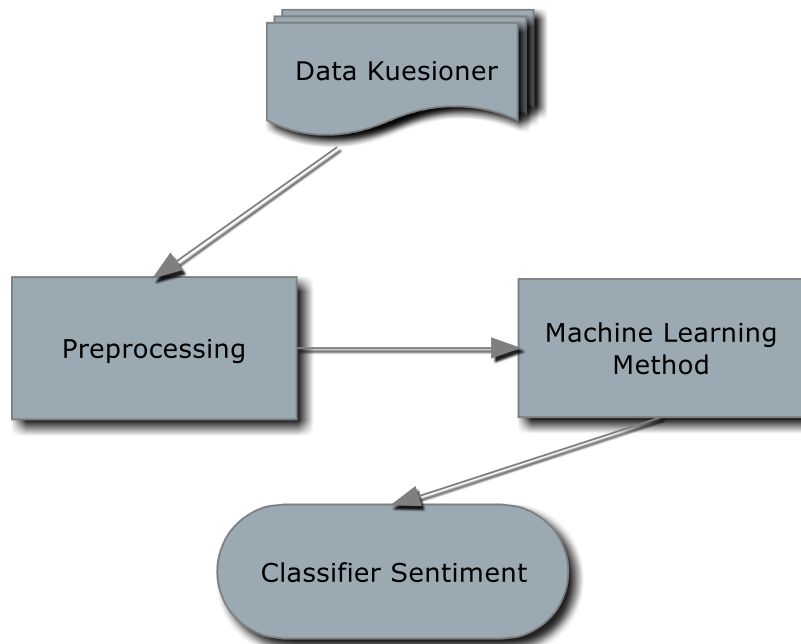
Keterangan :

Jawaban dari responden (mahasiswa) merupakan *data training*.

3. Data Output

Data ouput yang diharapkan adalah berupa *opini mining* dari jawaban responden yang mengacu pada data bobot jawaban yang disimpan pada database.

B. Block Diagram

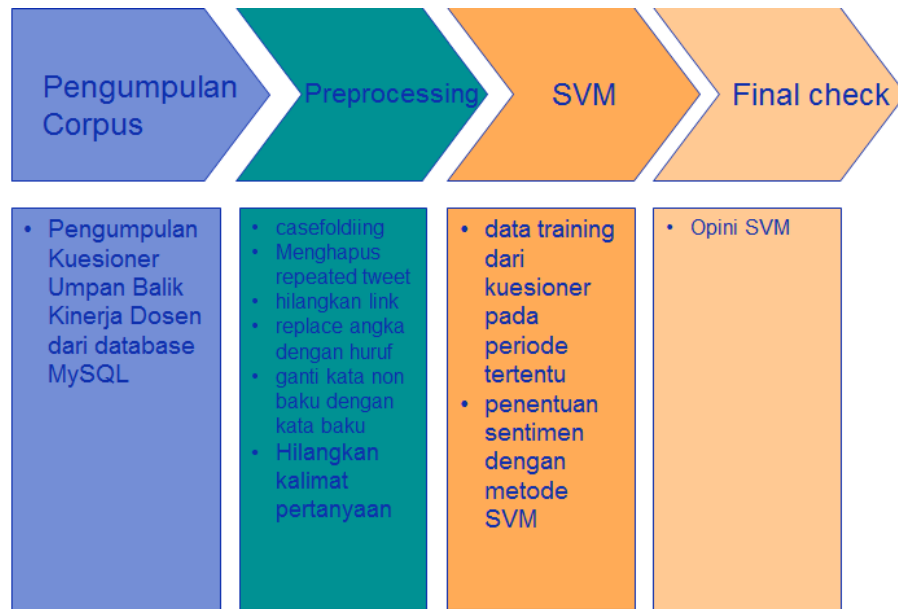


Gambar 4. Gambaran Umum Sistem yang akan dibangun

1. Perancangan

Pada tahap ini akan dilakukan perancangan *Metode Support Vector Machine (SVM)* yang akan diimplementasikan untuk *Opinion mining*. Data-data yang diambil dalam penelitian ini diambil dari evaluasi umpan balik kinerja dosen dengan memanfaatkan kuesioner online mahasiswa yang sudah disediakan, selanjutnya data tersebut disimpan dalam basis data. Langkah pertama dalam proses besar penelitian ini adalah melakukan *preprocessing* data kuesioner. Setelah dilakukan *preprocessing*, sistem akan mencari kalimat yang mengandung frasa opini, sehingga kalimat yang tidak mengandung frasa opini tidak diikutkan dalam proses selanjutnya. Kemudian sistem akan menentukan sentimen data kuesioner tersebut menggunakan *Metode Support Vector Machine*

(SVM) yang menghasilkan sentimen positif, sangat positif, negatif atau sangat negatif serta sentimen netral yang akan dipakai menjadi kesimpulan akhir.



Gambar 5. Proses Perancangan



Gambar 6. Diagram Alir Klasifikasi dengan SVM

2. Proses Klasifikasi Dengan Metode SVM

Proses klasifikasi menggunakan *SVM* dimulai dengan mengubah text menjadi data *vector*. *Vector* dalam penelitian ini memiliki dua komponen yaitu dimensi (*word id*) dan bobot. Bobot ini sering dikombinasikan ke dalam sebuah nilai *tf-idf*, secara sederhana dengan mengalikan mereka bersama-sama. Ada banyak variasi pada gagasan dasar *tf-idf*, tetapi implementasi langsung akan terlihat seperti:

```

<?php
$tfidf = $term_frequency * // tf
log( $total_document_count / $documents_with_term, 2); //idf
?>

```

idf adalah jumlah total dokumen atas hitungan yang berisi istilah tersebut. Jadi, jika ada 50 dokumen dalam koleksi, dan dua di antaranya terdapat istilah yang

menjadi pertanyaan, IDF akan menjadi $50 / 2 = 25$. Untuk menjadi akurat, harus memasukkan *query* dalam perhitungan IDF, jadi jika dalam koleksi ada 50 dokumen, dan 2 berisi istilah dari *query*, perhitungan yang sebenarnya akan $(50 + 1) / (2 + 1) = 51 / 3$. Diambil log dari IDF untuk memberikan beberapa penghalusan. Jika sebuah istilah A direpresentasikan dalam x buah dokumen, dan istilah B sejumlah 2x kali, maka istilah A adalah istilah yang lebih spesifik yang harus memberikan hasil yang lebih baik, tetapi belum tentu dua kali lebih baik. Kelembutan dari log adalah pemecahan perbedaan-perbedaan ini. Dokumen dapat dinyatakan sebagai list dari *term*. Sebuah contoh mapping dokumen untuk term (istilah) dalam bahasa PHP adalah sebagai berikut :

```
<?php
function getIndex() {
    $collection = array(
        1 => 'this string is a short string but a good string',
        2 => 'this one isn\'t quite like the rest but is here',
        3 => 'this is a different short string that\' not as short'
    );
    $dictionary = array();
    $docCount = array();
    foreach($collection as $docID => $doc) {
        $terms = explode(' ', $doc);
        $docCount[$docID] = count($terms);
        foreach($terms as $term) {
            if(!isset($dictionary[$term])) {
                $dictionary[$term] = array('df' => 0, 'postings' =>
                    array());
            }
            if(!isset($dictionary[$term]['postings'][$docID])) {
                $dictionary[$term]['df']++;
                $dictionary[$term]['postings'][$docID] = array('tf' => 0);
            }
            $dictionary[$term]['postings'][$docID]['tf']++;
        }
    }
    return array('docCount' => $docCount, 'dictionary' =>
        dictionary);
}
?>8
```

Kemudian menormalisasi tiap komponen dengan panjang dari *vector* sehingga bobot tersebut dinyatakan dalam 1 unit panjang. Masalah klasifikasi adalah sesuatu yang telah kita bahas sebelumnya, tetapi pada umumnya adalah tentang belajar yang memisahkan dua set contoh, dan berdasarkan hal tersebut menempatkan dengan benar contoh-contoh yang tak terlihat ke salah satu himpunan. Contohnya bisa berupa *filter spam*, di mana, diberikan pelatihan

⁸ Barber, I. 2009. *Support Vector Machines In PHP*

himpunan *mail spam* dan *non-spam* diharapkan untuk mengklasifikasikan email sebagai spam atau bukan spam.

SVM adalah sistem untuk melakukan hal itu, tetapi mereka hanya peduli tentang titik dalam ruang, daripada email atau dokumen. Untuk tujuan ini model ruang vektor digunakan untuk memberikan setiap kata dalam dokumen sebuah ID (dimensi) dan sebuah bobot berdasarkan seberapa penting keberadaannya dalam dokumen (posisi dokumen dalam dimensi itu). *SVM* mencoba untuk menemukan garis yang terbaik membagi dua kelas, dan kemudian mengklasifikasikan dokumen uji berdasarkan di sisi mana dari garis tersebut mereka muncul.

Format data input untuk klasifikasi *SVM* dalam penelitian ini adalah :

+1 1:0.049 45:0.0294....

Dengan masukan yang pertama +1 atau -1 menyatakan dua kelas (atau 0 untuk data yang akan diklasifikasi). Angka kedua menyatakan dimensi (*row id*) dan angka ketiga (setelah karakter “:”) menyatakan bobot dari term tersebut, tiap term dalam sebuah dokumen dipisahkan dengan spasi.

Intuisi yang mendorong *SVM* sebagai garis terbaik yang memisahkan kedua kelas adalah yang memiliki margin terbesar diantaranya dan contoh titik pelatihan terdekat di kedua sisinya. Oleh karena itu, *vektor* contoh penting adalah *vektor* yang menentukan margin tersebut - yang paling dekat dengan *dividing lines*. Ini adalah *support vector*, dan merupakan kombinasi dari *vector-vector* yang memberikan keputusan fungsi (kelas atau bukan kelas) untuk *classifier SVM*.

Fungsi klasifikasi dalam contoh kode adalah sebagai berikut:

```
<?php
protected function classify($rowID) {
    $score = 0;
    foreach($this->lagrangeMults as $key => $value) {
        if($value > 0) {
            $score += $value * $this->targets[$key] * $this->kernel($rowID, $key);
        }
    }
    return $score - $this->bias;
}
?>
```

Penilaian kemudian dibuat dengan menilai *score* apakah positif, apakah sangat positif, apakah negatif, apakah sangat negatif yang merepresentasikan di sisi mana dari garis pemisah dokumen berada. Sejauh ini fungsi kernel dapat diasumsikan sebagai *dot product* antara dua vector. Namun, *dot product* dalam

fungsi klasifikasi SVM tidak harus sebagai sebuah *dot product*, yang memungkinkan kita untuk memasang fungsi dengan tipe yang berbeda. Fungsi ini dapat menentukan secara efektif sebuah pemetaan dimensi pada ruang lingkup permasalahan, dan beberapa ruang lingkup dimensional yang lebih tinggi dimana kemungkinan data akan dipisahkan secara linier. Trik ini memungkinkan hasil klasifikasi yang baik pada sumber berbeda dan bervariasi, meskipun tidak mendukung beban dalam hal kompleksitas tambahan dan biaya komputasi.


3. Implementasi

Pada tahap ini akan dilakukan pembuatan kode program sampai proses pembangunan basis data.

4. Pengujian

Pada tahap ini, dilakukan perhitungan akurasi dari implementasi yang telah dilakukan. Metode yang digunakan adalah *precision*, *recall* dan *F-score*. Menurut Risbergen (1979), *Precision* adalah perbandingan antara jumlah kalimat yang mengandung opini baik sentimen positif, sangat positif, negatif maupun sangat negatif yang dapat dikelompokkan oleh sistem dengan jumlah seluruh kalimat yang diambil oleh system. Sedangkan *Recall* adalah perbandingan antara jumlah kalimat yang mengandung opini, baik sentimen positif, sangat positif, negatif maupun sangat negatif yang dapat dikelompokkan oleh sistem dengan jumlah kalimat yang mengandung opini, baik positif, sangat positif, negatif maupun sangat negatif yang berada dalam koleksi dokumen. *F-score* adalah 2 kali perbandingan antara perkalian nilai *Precision* dan nilai *Recall* dengan penjumlahan antara *Precision* dan *Recall*.

C. Desain Output
Tampilan Output

	Form Hasil Kuesioner Umpan Balik Mahasiswa <u>Ta 2012/2013</u> Program Studi <u>Manajemen Informatika</u> Jurusan Administrasi Bisnis Politeknik Negeri Banjarmasin		
	NIP : <input type="text"/> Nama : <input type="text"/> Pangkat/Golongan : <input type="text"/> Jabatan : <input type="text"/> Unit Kerja : <input type="text"/>	Nama Matakuliah XXXXXXXXXXXX	Kelas XXXX

No.	Keterangan	Opini
1	Point 1	Positif
2	Point 2	Positif
3	Point 3	Sangat Positif
4	Point 4	Positif
5	Point 5	Positif
6	Point 6	Negatif
7	Point 7	Positif
8	Point 8	Sangat Negatif
9	Point 9	Sangat Positif
10	Point 10	Netral
11	Point 11	Positif
12	Point 12	Negatif
13	Point 13	Negatif
14	Point 14	Positif

Keterangan Opini akan diperoleh setelah menekan tombol Proses.

VII. Batasan Masalah

1. Metode yang digunakan untuk membangun penilaian evaluasi umpan balik ini menggunakan *text mining* dengan metode *Metode Support Vector Machine (SVM)*.

2. Penelitian terbatas pada pengolahan data evaluasi umpan balik kinerja dosen dengan menentukan bobot opini dari suatu data kuesioner mahasiswa.
3. Bahasa Pemrograman yang digunakan adalah *PHP* dan *Borland C++*.
4. Database yang digunakan adalah *MySQL Server*.
5. *LimeSurvey*, aplikasi pengolah kuesioner.
6. Editor *Adobe Dreamweaver CS 6.6*

VIII. Metedologi Penelitian

1. Studi Literatur
Dengan mempelajari buku-buku referensi dan jurnal yang berkaitan dengan permasalahan penelitian yang diangkat serta mencari solusi yang terbaik. Topik bahasan utama yang dibutuhkan diantaranya adalah *text mining* dan *Metode Support Vector Machine (SVM)*.
2. Analisa
Melakukan uji coba secara teoritis terhadap masalah yang diangkat guna menganalisa apakah rancangan algoritma yang digunakan dapat menghasilkan solusi yang sesuai dengan tujuan penelitian.
3. Implementasi
Membuat program dari hasil rancangan algoritma yang telah dibuat untuk mengimplementasikan serta membuktikan bahwa hasil analisa secara teoritis yang telah dilakukan benar-benar sesuai yang diharapkan.
4. Pengujian
Pengujian dilakukan untuk melihat apakah Data yang telah menjadi input akan diproses sesuai dengan output yang harapan. Hal ini juga dilakukan untuk mengevaluasi apakah metode yang diusulkan mampu menjawab tujuan yang telah usulkan.
5. Dokumentasi
Merupakan langkah akhir, penyusunan laporan mulai dari latar belakang permasalahan sampai dengan pengambilan kesimpulan akan dijelaskan dalam tahap dokumentasi ini.

IX. Jadwal Penelitian

Penelitian ini akan dijadwalkan menjadi beberapa detail kegiatan. Masing-masing detail dari kegiatan yang dilakukan akan dijabarkan menjadi beberapa tabel yang memuat detail waktu pengerjaan dari tesis ini.

No	Kegiatan	Bulan					
		1	2	3	4	5	6
1	Studi Literatur						
2	Analisa						
3	Implementasi						
4	Pengujian						
5	Dokumentasi						

X. Daftar Pustaka

- Anonym. 2010. *Sentiment Analysis*. [Online]. Tersedia di:
http://en.wikipedia.org/wiki/Sentiment_Analysis.
- Anonym. 2010. *Text Mining*. [Online]. Tersedia di:
http://en.wikipedia.org/wiki/Text_mining.
- Anonym. 2010. *Support Vector Machine* [Online]. Tersedia di:
http://en.wikipedia.org/wiki/Support_vector_machine.
- Barber, I. 2009. *Simple Search : The Vector Space Model*. [Online]. Tersedia di:
<http://phpir.com/simple-search-the-vector-space-model>
- Barber, I. 2009. *Support Vector Machines In PHP*. [Online]. Tersedia di:
<http://phpir.com/support-vector-machines-in-php>
- Bing Liu, *Opinion Mining*. Department of Computer Science University of Illinois at Chicago 851 S. Morgan Street Chicago, IL 60607-0753.
Tersedia di: <http://www.cs.uic.edu/~liub/FBS/opinion-mining.pdf>
- Colas, F. & Brazdil, P. 2005. *Comparison of SVM and Some Older Classification Algorithms in Text Classification Tasks*.
- Cortes, C. & Vapnik, V. 1995. *Support-Vector Networks*. *Machine Learning*, 20. [Online]. Tersedia di:
<http://www.springerlink.com/content/k238jx04hm87j80g/>.
- Dehaff, M. 2010. *Sentiment Analysis, Hard But Worth It!*. [Online]. Tersedia di:
http://www.customerthink.com/blog/sentiment_analysis_hard_but_worth_it
- Jason D. M. Rennie & Ryan Rifkin. *Improving Multiclass Text Classification with the Support Vector Machine*. [Online]. Tersedia di:
<http://www.ai.mit.edu>

- Jenkins, M. C. 2011. *How Sentiment Analysis works in machines*. [Kuliah]. Tersedia di: <http://www.slideshare.net/mcjenkins/how-sentiment-analysisworks/> download
- Mihalcea, R. , Banea, C. & Wiebe, J. 2007. *Learning Multilingual Subjective Language via Cross-Lingual Projections*. *Proceedings of the Association for Computational Linguistics (ACL)*. pp. 976–983. [Online]. Tersedia di: <http://www.cse.unt.edu/~rada/papers/mihalcea.acl07.pdf>.
- Pang, B. & Lee, L. 2004. *A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts*. *Proceedings of the Association for Computational Linguistics (ACL)*. pp. 271–278. [Online]. Tersedia di: <http://www.cs.cornell.edu/home/llee/papers/cutsent.home.html>.
- Pang, B. & Lee, L. 2005. *Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales*. *Proceedings of the Association for Computational Linguistics (ACL)*. pp. 115–124. [Online]. Tersedia di: <http://www.cs.cornell.edu/home/llee/papers/pang-lee-stars.home.html>
- Pang, B. & Lee, L. 2008. *Subjectivity Detection and Opinion Identification*. *Opinion Mining and Sentiment Analysis*. Now Publishers Inc. [Online]. Tersedia di: <http://www.cs.cornell.edu/home/llee/opinion-mining-sentiment-Analysissurvey.html>.
- Peter Cabena, (1998), *Discovering Data Mining: From Concept to Implementation*, Chapter 3, Prentice Hall
- Snyder B. & Barzilay R. 2007. *Multiple Aspect Ranking using the Good Grief Algorithm*. *Proceedings of the Joint Human Language Technology/North American Chapter of the ACL Conference (HLT-NAACL)*. pp. 300–307. [online]. Tersedia di: http://people.csail.mit.edu/regina/my_papers/ggranker.ps.
- Su, F. & Markert, K. 2008. *From Words to Senses: a Case Study in Subjectivity Recognition*. *Proceedings of Coling 2008, Manchester, UK*. [Online]. Tersedia di: <http://www.comp.leeds.ac.uk/markert/Papers/Coling2008.pdf>.
- Sundani, D. 2008. *Perangkat lunak SPSS sebagai alat untuk analisa hubungan kinerja dosen dengan keberhasilan belajar*. *Jurnal Informatika Komputer*, vol.13.
- Tan, P. N., Steinbach, M. & Kumar, V. 2006. *Introduction to Data Mining*. Boston : Pearson Addison Wesley.
- Undang-Undang RI No.14 tentang Guru dan Dosen,(2005).

- Yuki Uchida, dkk, (2009), *Extract of Important Keywords in Free Text of Questionnaire Data and Visualition of Relationship Among Sentences*. Korea, August 20-24, FUZZ-IEEE
- Wibisono, Y. 2005. *Clustering Berita Berbahasa Indonesia*. [Online]. Tersedia di: http://fpmipa.upi.edu/staff/yudi/KNSI_Clustering_yudi_masayu.pdf
- Widodo, Pudjo Prabowo, dkk. 2013. *Penerapan Data Mining dengan Matlab*. Rekayasa Sains. Bandung.
- Wulandini, F. & Nugroho, A. N. 2009. *Text Classification Using Support Vector Machine for Webmining Based Spation Temporal Analysis of the Spread ofTropical Diseases*. *International Conference on Rural Information and Communication Technology 2009*.

XI. Daftar Riwayat Hidup

Nama : Tajudin Noor, S.Kom
 Tempat, Tgl.Lahir : Kandangan, 11 Mei 1973
 Pekerjaan : Pegawai Negeri Sipil (PNS)
 Riwayat Pendidikan : S-1 Sistem Informasi pada STMIK Indonesia Banjarmasin
 Riwayat Pekerjaan :

No.	Masa Kerja	Pangkat / Gol.	Kantor
1.	2006 – Sekarang	Penata Muda TK. I/ III-b	Politeknik Negeri Banjarmasin Jurusan Administrasi Bisnis

Revisi Proposal tesis

1. Harus ada ranking hasil *opinion* (*sentiment strength detection*)

Jawaban :

Pada bagian ruang lingkup telah dilakukan koreksi pada hasil opini yang semula menghasilkan opini positif dan opini negatif saja menjadi opini positif, sangat positif, negatif, dan sangat negatif. Dimana opini positif memiliki *score* antara 2 dan 4, *score* 5 untuk opini sangat positif, sedangkan opini negatif memiliki *score* antara -2 dan -4, untuk opini sangat negatif memiliki *score* -5. Bila merupakan bukan opini maka diberi nilai *score* 1 atau netral.

Pada halaman 17 pada gambar 6 yang menggambarkan diagram alir klasifikasi dengan SVM dilakukan modifikasi menjadi sebagai berikut :

