

1. Pendahuluan

1.1 Latar Belakang Masalah

Kemajuan dunia teknologi informasi, mendorong pesatnya perkembangan teknologi internet. Layanan aplikasi pada internet seperti blog, jejaring sosial, dan forum online sudah merupakan aktivitas keseharian yang lumrah. Layanan aplikasi dalam internet tersebut menghasilkan banyak sekali informasi yang dapat kita kategorikan sebagai opini dan fakta. Fakta merupakan segala sesuatu yang tertangkap oleh indra manusia atau data keadaan nyata yang terbukti dan telah menjadi suatu kenyataan. Sedangkan opini adalah pendapat, ide atau pikiran untuk menjelaskan kecenderungan atau preferensi tertentu terhadap perspektif dan ideology akan tetapi bersifat tidak objektif [8].

Opini yang banyak terdapat dalam media online dapat di analisis dan diklasifikasikan menjadi dua bagian, yaitu opini positif dan opini negatif, analisis ini disebut sebagai analisis sentimen. Beberapa opini yang terdapat pada internet merupakan suatu pandangan atau *review* terhadap sebuah produk ataupun layanan tertentu. Beberapa opini pelanggan terhadap suatu produk ini sangat bermanfaat untuk perusahaan tersebut ataupun pelanggan lainnya. Analisis dalam opini penting dilakukan karena bermanfaat terhadap identifikasi produk atau kecenderungan pasar bagi perusahaan, sehingga dapat diambil langkah-langkah strategis berikutnya. Besarnya pengaruh dan manfaat dari analisis sentimen, menyebabkan penelitian ataupun aplikasi mengenai analisis sentimen berkembang pesat, bahkan di Amerika kurang lebih 20-30 perusahaan yang memfokuskan pada layanan analisis sentiment [11]. Proses analisis sentiment untuk opini berbahasa Indonesia juga mulai berkembang untuk tujuan penelitian dan untuk pengambilan langkah strategis. Mengingat bahwa Bahasa Indonesia merupakan bahasa yang jamak dan variasi penggunaan Bahasa Indonesia dalam internet sering menggunakan kata yang tidak baku dan membutuhkan penanganan yang berbeda dibanding bahasa Inggris, maka faktor-faktor tersebut mendorong perlunya dilakukan opinion mining berbahasa Indonesia.

Banyak metode *machine learning* yang digunakan dalam proses data mining sekaligus analisis sentiment. Salah satu metode klasifikasi *machine learning* yang terkemuka adalah *Support Vector Machine (SVM)*. Pada dasarnya metode *SVM* bekerja dengan menemukan *hyperplane* terbaik yang memisahkan dua *class* pada input space. *SVM* memiliki performansi yang lebih baik dibandingkan dengan metode *machine learning* lainnya. Penggunaan kernel dalam metode *Support Vector Machine* belakangan ini mendapat cukup banyak perhatian karena dapat diaplikasikan dalam berbagai bidang penelitian. Dalam penelitian ini telah dilakukan analisis perbandingan beberapa kernel dalam *SVM*. Kernel dalam *SVM* yang umum dan sering digunakan adalah kernel *Linear*, dimana kernel *Linear* mampu melakukan klasifikasi data yang terpisah secara linear. *Support Vector Machine* dengan menggunakan kernel *Linear* telah banyak digunakan dan terbukti memberikan hasil akurasi yang optimal dalam permasalahan linear. Namun pada umumnya, masalah dalam dunia nyata sulit dilakukan pemisahan secara linear,

dalam arti lebih banyak yang bersifat non linear. Dalam aspek sentiment, beberapa opini tidak sepenuhnya dapat dipisahkan menggunakan kernel *Linear*, untuk mengatasi masalah non-linear, kernel yang dapat digunakan adalah Kernel Gaussian atau *Radial Basic Function (RBF)*. Dimana kernel ini merupakan kernel yang mampu merefleksikan informasi non linear[10]. Sehingga kernel lain yang akan dibandingkan dalam penelitian ini adalah kernel *Radial Basic Function* atau biasa disebut *Gaussian Kernel*. Dalam representasi data masukan untuk *SVM* kedua kernel ini menggunakan kamus leksikal yang disebut sentiwordnet bahasa Indonesia. Kemudian hasil klasifikasi opini menggunakan kedua kernel tersebut akan dianalisis dan dibandingkan akurasi dan performansinya.

1.2 Perumusan Masalah

Berdasarkan pada latar belakang di atas, permasalahan yang akan diuraikan dan diteliti adalah:

1. Bagaimana cara menerapkan metode *Support Vector Machine (SVM)* dengan kernel *Linear* dan *Radial Basic Function* untuk proses klasifikasi opini.
2. Bagaimana performansi antara dua dataset opini yang berbeda terhadap *Support Vector Machine (SVM)* dengan kernel *Linear* dan *Radial Basic Function*.
3. Bagaimana analisis parameter yang berpengaruh pada metode *Support Vector Machine (SVM)* yaitu parameter C dalam kernel *Linear* dan kombinasi parameter C dan γ dalam kernel *Radial Basic Function* pada proses klasifikasi.

1.3 Batasan Masalah

Dalam pengerjaan Tugas Akhir ini, ada beberapa hal yang dibatasi yaitu sebagai berikut:

1. Opini yang dianalisis terbatas pada opini suatu layanan pertamina dalam blog pastipas pertamina dan produk handphone dalam halaman web tabloidpulsia.
2. Opini yang akan diretrieve harus berbahasa Indonesia
3. Opini yang akan mengalami klasifikasi tidak menangani kata yang disingkat dan tidak lengkap. Namun untuk kata yang tidak baku, dilakukan proses cleaning pada tahap *preprocessing*
4. Tidak menangani opini yang bermakna ambigu.
5. Proses klasifikasi hanya terbatas pada opini positif dan negatif
6. Menggunakan sentiwordnet berbahasa Indonesia
7. Menggunakan *LibSVM* sebagai library untuk membangun *SVM* baik dengan kernel *Linear* maupun *Radial Basic Function*
8. Data yang akan di ambil hanya berbentuk teks murni ataupun plain teks tidak termasuk suara, image atau video.
9. Data yang diproses dianalisis secara offline. Yaitu sistem tidak terhubung pada internet.

1.4 Tujuan

Adapun tujuan yang ingin dicapai dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

1. Menerapkan metode *Support Vector Machine (SVM)* dengan kernel *Linear* dan *Radial Basic Function* untuk proses klasifikasi opini.
2. Menganalisis performansi antara dua dataset opini yang berbeda terhadap *Support Vector Machine (SVM)* dengan kernel *Linear* dan *Radial Basic Function*.
3. Menganalisis parameter yang berpengaruh pada metode *Support Vector Machine (SVM)* yaitu parameter C dalam kernel *Linear* dan kombinasi parameter C dan γ dalam kernel *Radial Basic Function* pada proses klasifikasi.

1.5 Metodologi Penyelesaian Masalah

1. Melakukan studi literatur
Mempelajari literatur-literatur mengenai *opinion mining* dan metode *Support Vector Machine* dengan mempelajari jenis-jenis kernel apa saja yang terdapat dalam *SVM*. Kernel yang akan digunakan dalam penelitian ini adalah kernel *Linear* dan *Radial Basic Function*.
2. Perancangan sistem dan pengumpulan data
Dalam tahap awal ini dilakukan perancangan sistem yang akan dibangun, mulai dari *preprocessing* termasuk tahap apa saja yang dilakukan dalam *preprocessing* dan pengumpulan dataset yang berupa dokumen yang diperoleh dari website berbahasa Indonesia www.tabloidpulsa.com sebanyak 500 opini, dan dataset opini yang berasal dari blog pasti pas Pertamina sebanyak 132 opini yaitu <http://pastipas.pertamina.com/pendapat.asp>. dimana dataset ini terdiri dari opini yang memiliki komposisi opini positif dan negatif yang sama jumlahnya.
3. *Preprocessing* dataset
Tahapan *preprocessing* bertujuan agar dataset siap digunakan menjadi inputan dalam proses klasifikasi menggunakan *Support Vector Machine* baik dengan kernel *Linear* maupun *Radial Basic Function*. Selain itu dilakukan proses pembobotan menggunakan kamus leksikal sentiwordnet berbahasa Indonesia. Dalam kamus ini setiap kata memiliki nilai dalam sentimen positif maupun negatif, dari nilai tersebut dijumlahkan dan menghasilkan data numerik yang berguna sebagai masukan sistem *SVM*.
4. Implementasi menggunakan metode *SVM*
Pada tahap ini, rancangan sistem yang telah dibuat diimplementasikan menggunakan bahasa pemrograman Java. Metode *Support Vector Machine* bekerja dengan cara menentukan *hyperplane* terbaik yang memisahkan dua buah *class* pada *input space*, yang akan menghasilkan opini positif dan opini negatif. Dalam tahap ini dilakukan klasifikasi dengan dua metode *SVM* menggunakan *Linear* kernel dan *Radial Basic Function* kernel.
5. Analisis dan pengujian
Pada tahap ini dilakukan beberapa analisis mengenai performansi dan akurasi metode *Support Vector Machine* dengan kernel *Linear* dan *RBF*, serta

parameter yang berpengaruh terhadap hasil klasifikasi opini positif dan negatif.

6. Pengambilan kesimpulan dan pembuatan laporan

Setelah dilakukan analisis, maka dapat diambil kesimpulan serta dapat segera melakukan pembuatan laporan Tugas Akhir sebagai dokumentasi kegiatan yang dilakukan.