

Abstrak

Twitter adalah salah satu sosial media yang banyak digunakan oleh orang Indonesia. Didalam *Twitter* banyak ditemukan beragam permasalahan atas suatu peristiwa yang banyak menimbulkan opini-opini atau pandangan dari masyarakat, salah satunya peristiwa atau *event* pemilihan Gubernur Jawa Barat 2018. Pemilihan Gubernur Jawa Barat 2018 ini diikuti oleh calon-calon dari tokoh politik yang berkompeten di bidangnya, yaitu Ridwan Kamil, Deddy Mizwar, Sudrajat dan TBHasanuddin. Dari suatu peristiwa atau *event* ini dapat dianalisis opini atau sentimennya baik berupa sentimen positif, negatif atau netral dengan cara menggunakan Analisis Sentimen. Dan data yang diambil bersumber dari kicauan-kicauan atau *tweet* mengenai calon-calon Gubernur di Pemilihan Gubernur Jawa Barat 2018 di *Twitter* dengan menggunakan *keyword* nama calonnya seperti “ridwankamil”, “deddymizwar”, “sudrajat”, dan “tbhasanuddin”, sementara untuk proses pengklasifikasian kelasnya, baik berupa kelas positif, negatif atau netral digunakan algoritma *Naïve Bayes Classifier* dan juga pembobotan TF-IDF. Lalu tujuan dari penelitian ini adalah untuk melihat arah dari pandangan masyarakat mengenai calon-calon Gubernur di Pemilihan Gubernur Jawa Barat 2018 dengan memperhatikan jumlah dari kelas positif, negatif dan netral per nama calon Gubernurnya. Dan untuk skenario yang digunakan di dalam penelitian ini adalah dengan penggunaan *stopword* dan tanpa *stopword* serta fitur N-gram seperti Unigram, Bigram, Trigram juga menggunakan komposisi data training sebesar 60%,70% dan 80% pada pengujian akurasi dan pengujian pengklasifikasian kelas sentimen. Hasil yang diperoleh di dalam penelitian ini untuk jumlah akurasi tertinggi pada penggunaan *stopword* yaitu sebesar 97.92% yang diperoleh dari data TBHasanuddin pada fitur Trigram dan komposisi data training 60%. Lalu untuk akurasi terkecil pada penggunaan *stopword* diperoleh dari data Deddy Mizwar dengan fitur Trigram dan persentase data training 60% dengan jumlah akurasinya sebesar 72.92%. Sementara untuk perhitungan akurasi pada penggunaan tanpa *stopword* pada akurasi tertinggi diperoleh dari data TBHasanuddin dengan fitur Trigram dan persentase data training 60% dengan jumlah akurasi sebesar 97.53%. Dan untuk akurasi terkecil pada penggunaan tanpa *stopword* diperoleh dari data Deddy Mizwar dengan fitur Bigram dan persentase data training 70% dengan jumlah akurasi sebesar 72.22%. Dan untuk hasil dari pengujian klasifikasi sentimen untuk kelas positif tertinggi diperoleh dari data data Sudrajat dengan fitur Trigram dan persentase data testing 40% dengan jumlah nilai kelas positif pada klasifikasinya adalah 104 dan akurasi sebesar 95%. Lalu untuk penilaian jumlah klasifikasi sentimen pada kelas netral tertinggi diperoleh dari data TBHasanuddin dengan fitur Trigram dan persentase data testing 40% dengan jumlah nilai kelas netralnya adalah 203 kelas dan akurasi sebesar 95.83%. Sementara untuk nilai kelas negatif dengan jumlah nilai kelas negatif tertinggi diperoleh dari data Deddy Mizwar dengan fitur Trigram dan persentase data testing 40% dengan jumlah nilai kelas negatif sebesar 91 kelas dan akurasi sebesar 77.5%.

Kata kunci : Twitter, Analisis Sentimen, Machine Learning, TF-IDF, Naïve Bayes, Pilgub Jabar 2018.