

LAPORAN TUGAS NLP PEKAN 10 – CLASSIFICATION TEXT

1. Algoritma yang digunakan adalah Decision Tree
2. Lakukan evaluasi dan analisis hasil klasifikasi berdasarkan Precision, Recall dan F-Measure

	Predicted 0	Predicted 1	Predicted 2	
Is 0	130	0	10	
Is 1	2	70	16	
Is 2	5	10	43	

	precision	recall	f1-score	support
0	0.95	0.93	0.94	140
1	0.88	0.80	0.83	88
2	0.62	0.74	0.68	58
accuracy			0.85	286
macro avg	0.82	0.82	0.82	286
weighted avg	0.86	0.85	0.85	286

Pertama, lihat terlebih dahulu pada table atas yang menyatakan bahwa berapa hasil prediksi yang sesuai dengan kenyataannya. Untuk hasil label dari prediksi label 0 memiliki hasil prediksi benar hanya 130/137. Sedangkan untuk hasil prediksi benar dari label 1 hanya 70 dari 80 data saja yang sesuai dengan data asli. Kemudian untuk hasil prediksi benar dari label 2 hanya ada 43 dari 69 data yang sesuai dengan data asli.

Precision adalah tingkat ketepatan antara informasi yang diminta (data set) dengan jawaban yang diberikan oleh system (y_{test}). Precision merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Oleh karena itu di setiap labelnya dihitung nilainya masing-masing precisionnya sehingga menghasilkan nilai pada table kedua.

Rumus : $TP / (TP+FP)$

Label 0 : $130 / (130+2+5) = 130/137 = 0.95$

Label 1 : $70 / (70+10) = 70/80 = 0.88$

Label 2 : $43 / (43+10+16) = 43/69 = 0.62$

Nilai pada label 1 lebih tinggi dibanding yang lain karena label 0 karena jika dihitung, semua sentiment yang diprediksi bernilai 0, 130 dari 137 nilai pada data aslinya memang benar 0, jadi oleh karena itu nilai precision untuk label 1 adalah 0.95 atau 95 %. Yang

bernilai tinggi setelahnya adalah label 1 karena hasil prediksi label 1 lebih banyak benarnya dibandingkan dengan hasil prediksi label 2.

Recall adalah tingkat keberhasilan system dalam menemukan kembali sebuah informasi. Recall Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif ditambah data yang salah atau data keseluruhan pada label tersebut. Oleh karena itu di setiap labelnya dihitung nilainya masing-masing recallnya sehingga menghasilkan nilai pada table kedua.

Rumus : $TP / (TP+FN)$

Label 0 : $130 / (130+10) = 130/140 = 0.93$

Label 1 : $70 / (70+2+16) = 70/88 = 0.80$

Label 2 : $43 / (43+5+10) = 43/58 = 0.74$

Pada hasil perhitungan tersebut sudah jelas terlihat nilai recall pada label 0 lah yang paling tinggi. Hal ini dikarenakan banyak data prediksi atau sekitar 93% yang berlabel 0 bernilai benar sesuai dengan dataset yang ada. Kemudian selanjutnya ada label 1, 80% persen dari prediksi yang dijatuhkan pada label 1 bernilai true. Kemudian nilai recall paling rendah ada pada label 2, yaitu hanya 74% prediksi label yang dijatuhkan pada label 2 memiliki nilai true atau benar sesuai dengan dataset.

F-Measure atau F1-Score adalah perhitungan evaluasi dalam informasi yang mengkombinasikan nilai dari recall dan precision.

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Rumus :

Label 0 : $2 \times ((0.95 \times 0.93) / (0.95 + 0.93)) = 0.94$

Label 1 : $2 \times ((0.88 \times 0.80) / (0.88 + 0.80)) = 0.83$

Label 2 : $2 \times ((0.62 \times 0.74) / (0.62 + 0.74)) = 0.68$

Nilai tertinggi dari F-Measure ada pada label 0 karena nilai precision dan recallnya juga merupakan nilai yang tertinggi..

3. Lakukan analisis terhadap kinerja algoritma klasifikasi berdasarkan :

a. Pengaruh persentase data latih dan data uji

Kinerja algoritma klasifikasi decision tree termasuk yang lumayan baik, karena jika dilihat dari presentase akurasi yaitu 85% itu termasuk ke dalam kategori baik karena sudah diatas 80%. Jika dilihat nilai presisi dan recallnya pun 66% nya berada di atas 80%, hanya ada 1 label saja yang nilainya dibawah 80%.

b. Waktu pelatihan yang diperlukan

Waktu pelatihan yang digunakan untuk membangun modelnya adalah 754 ms, dan waktu yang digunakan untuk menguji model tersebut adalah 330ms, dua kali lebih cepat dibandingkan dengan membangun model tersebut.

c. Data yang salah diklasifikasikan, jelaskan mengapa data tersebut salah klasifikasi!

Lampirkan juga data yang salah diklasifikasikan.

Terjadi kesalahan pada saat pengklasifikasian data uji bisa disebabkan oleh data latih yang kurang banyak modelnya, jadi kurang banyak vocab yang ada pada data model tersebut. Hal tersebut dapat menyebabkan salah pengklasifikasian saat kata yang ada pada data uji tidak ada pada data latih.

Note : Hasil prediksi, nilai recall, precision, f-measure dan akurasi dapat berubah-ubah.