

DAFTAR ISI

COVER	i
PERNYATAAN	ii
PRAKATA	ii
ARTI LAMBANG DAN SINGKATAN	vi
ABSTRACT	vii
INTISARI	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan masalah	3
1.3 Keaslian penelitian	3
1.4 Tujuan Penelitian	5
1.5 Batasan Masalah	5
1.6 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Landasan Teori	10
2.2.1 SMS Spam	10
2.2.2 Klasifikasi Teks	11
2.2.3 <i>Term Frequency – Inverse Document Frequency</i> (TF-IDF)	13
2.2.4 Metode <i>Principal Component Analysis</i> (PCA)	15
2.2.5 Metode Klasifikasi <i>Naïve Bayes Classifier</i> (NBC)	21

2.2.6	Korelasi	25
2.2.7	Metode Klasifikasi <i>Correlated-Naïve Bayes Classifier</i> (C-NBC) ..	27
2.3	Hipotesis	31
BAB III METODOLOGI.....		32
3.1	Alat dan Bahan Penelitian.....	32
3.1.1	Alat Penelitian	32
3.1.2	Bahan Penelitian.....	32
3.2	Alur Penelitian	33
3.2.1	Pengumpulan <i>Dataset</i>	34
3.2.2	<i>Preprocessing</i>	34
3.2.3	Pembobotan Fitur	38
3.2.4	Seleksi Fitur	40
3.2.5	Klasifikasi	43
3.2.6	Evaluasi	44
3.3	Rancangan Percobaan.....	45
3.3.1	Algoritme NBC dan CNBC tanpa Metode PCA.....	45
3.3.2	Algoritme NBC dan CNBC menggunakan Metode PCA	46
3.4	Cara Analisis.....	48
3.3.1	Uji Normalitas	48
3.3.2	Uji Hipotesis	50
BAB IV HASIL DAN PEMBAHASAN.....		52
4.1	Hasil <i>Preprocessing</i>	52
4.2	Hasil Pembobotan Fitur	53
4.3	Hasil Seleksi Fitur.....	54
4.4	Evaluasi Percobaan Metode Klasifikasi.....	56
4.5	Uji Normalitas	61
4.6	Uji Hipotesis.....	63
BAB V KESIMPULAN DAN SARAN.....		65
5.1	Kesimpulan	65
5.2	Saran	65



DAFTAR PUSTAKA	67
LAMPIRAN	1