

BAB II

TINJAUAN PUSTAKA

Penelitian mengenai implementasi analisis sentimen dengan menggunakan metode *Naïve Bayes* bukanlah yang pertama kali dilakukan. Telah ada penelitian-penelitian sebelumnya. Oleh karena itu, penelitian akan meneruskan apa yang telah dirintis para penelitian sebelumnya. Akan tetapi, penelitian yang penulis lakukan memiliki objek dan studi kasus yang berbeda dengan penelitian sebelumnya agar memperbaiki penelitian sebelumnya. Pada penelitian ini, penulis melakukan kajian ulang terhadap penelitian terdahulu dengan studi kasus yang berbeda agar dapat memperbaiki penelitian yang sudah dilakukan sebelumnya. Berikut merupakan penelitian terdahulu yang dapat digunakan sebagai acuan penulis:

Penelitian yang pertama yang dilakukan oleh (Afshoh, 2017) terkait persepsi masyarakat terhadap kenaikan harga jual rokok. Data yang digunakan dalam penelitian ini berupa *text* dalam media sosial yaitu *twitter*, terkait opini positif, negatif, maupun netral terhadap kenaikan harga jual rokok dengan jumlah data train 150 positif, 150 negatif, dan 50 netral. Hasil yang didapatkan yaitu bahwa opini positif lebih banyak terbentuk dalam menanggapi berita tersebut.

Penelitian yang dilakukan oleh (Santosa, 2016) terkait Situs *Web* Jawa Tengah dan diklasifikasikan menggunakan metode *Naive Bayes Classification* (NBC). Dengan dilakukannya penelitian ini bertujuan untuk dapat mengidentifikasi konten *e-government*. Hasil dari penelitian diatas yaitu dengan menggunakan data *training* sebanyak 160 dokumen politik dan 222 dokumen ekonomi dan menggunakan 40 data *testing* menunjukkan nilai akurasi yang baik yaitu sebesar 85%.

Penelitian menurut (Kurniawan, 2017) terkait berita *twitter* menggunakan metode *Improved Naive Bayes* untuk mengkategorikan *tweets* berdasarkan isi berita. Pada penelitian ini menggunakan delapan kategori berita yaitu: ekonomi,

entertainment, olahraga, teknologi, kesehatan, makanan, otomotif, dan travel dengan hasil nilai *precision* 0.96, *recall* terbaik pada data *training* dengan nilai 0.78 dan *f-measure* sebesar 0.86. Jumlah data *training* mempengaruhi performa, karena semakin banyak data *training* yang digunakan maka semakin baik.

Penelitian yang dilakukan oleh (Ramadhani, 2015) terkait *text mining* dengan menggunakan metode *Support Vector Machine* (SVM) dan *Naive Bayes Classifier* (NBC) dalam pengklasifikasian terkait perhitungan akurasi. Diperoleh bahwa tingkat akurasi tertinggi yaitu pada metode NBC sebesar 62,60%. Data hasil klasifikasi kemudian dihitung tingkat akurasinya menggunakan fungsi kernel dan proporsi data *training* dan *testing*. Jumlah data *training* 5412 dan data *testing* 2701 hasil klasifikasi, diperoleh hasil untuk kelas positif sebesar 44,50% dan kelas negatif sebesar 55,50%.

Dalam penelitiannya (Suryono, Utami, & Lufthi, 2018) terkait klasifikasi sentimen pada media sosial *twitter* dengan NBC (*Naive Bayes Classifier*). Klasifikasi sentimen digunakan untuk mencari kata-kata yang bermakna positif, negatif dan netral melalui *tweet* pada *twitter*. Data yang diperoleh kemudian diklasifikasikan menggunakan metode *naive bayes* dan akan dihitung nilai akurasinya. Didapatkan hasil uji coba sebanyak 3 kali, nilai akurasi yang pertama sebesar 64,95%, yang kedua sebesar 66,36% dan yang ketiga sebesar 66,79%. Dari proses klasifikasi didapatkan opini positif sebesar 28%, opini negatif sebesar 20%, dan opini netral sebesar 52%. Berdasarkan hasil sentimen yang didapatkan, opini netral yang paling banyak apabila dikaitkan dengan topik Presiden Joko Widodo dan pemerintahannya.

Penelitian yang dilakukan oleh (Falahah & Nur, 2015) yang bertujuan untuk membangun aplikasi klasifikasi opini dengan pendekatan *Naive Bayes* untuk mengklasifikasikan kata-kata. Sumber *text* diperoleh dari akun *twitter* Dinas Pemerintah Kota Bandung. Hasil dari analisis tersebut yaitu menghasilkan nilai akurasi sebesar 73% dengan jumlah data *training* 100 negatif, 100 positif, dan 100 netral.

Dalam penelitiannya (Hamzah, 2012) yaitu mengklasifikasikan teks berita dan abstrak akademis dari berbagai jurusan menggunakan metode *Naive Bayes*

Classifier. Jumlah data yang digunakan sebanyak 1000 dokumen berita dan 450 dokumen abstrak. Hasil pengujian menunjukkan bahwa akurasi dokumen berita sebesar 91%, sedangkan pada dokumen abstrak nilai akurasinya sebesar 82%.

Penelitian yang dilakukan oleh (Faishal Nuruz Zuhri, 2017) bertujuan untuk melihat sentimen masyarakat mengenai entitas tertentu. Sentimen dapat di ekspresikan dalam bentuk positif atau negatif. Penelitian ini mengklasifikasikan sentimen masyarakat terhadap *brand* Smartfren, bagaimana pandangan masyarakat terkait *brand* Smartfren dapat tertuang dalam sentimen baik itu positif maupun negatif. Penelitian ini mengambil komentar-komentar *user* terhadap *brand* Smartfren di forum Kaskus. Hasil dari penelitian ini yaitu dari data 6338 data uji, 4049 berhasil terklasifikasikan kedalam sentimen positif dan 3233 data bersentimen negatif. Dengan nilai akurasi sebesar 98,40%. Hasil sentimen positif berupa kata-kata smartfren, paket, *true unlimited*, *speed*, lancar, dan kencang, untuk kata-kata yang bersentimen negatif antara lain yaitu paket, smartfren, beli, *true unlimited*, lamban, masalah, dan gangguan.

Tabel 2. 1 Penelitian Sebelumnya

Penulis	Judul Penelitian	Metode Penelitian	Hasil
(Afshoh, 2017)	Analisa Sentimen menggunakan <i>Naive bayes</i> untuk Melihat Persepsi Masyarakat terhadap Kenaikan Harga Jual Rokok pada Media Sosial <i>Twitter</i>	Analisa Sentimen dan <i>Naive bayes classifier</i>	Sentimen positif paling banyak terbentuk untuk menanggapi wacana kenaikan harga jual rokok dan hasil pengujian kinerja sistem menggunakan data <i>training</i> 150 positif, 150 negatif, dan 50 netral yang menunjukkan hasil klasifikasi dokumen yang lebih baik.
(Santosa, 2016)	<i>Naive Bayes Classification</i> pada Klasifikasi Dokumen untuk Identifikasi Konten <i>E-Government</i>	<i>Naive Bayes Classifier</i>	Hasil dari penelitian diatas yaitu dengan menggunakan data <i>training</i> sebanyak 160 dokumen politik dan 222 dokumen ekonomi dan menggunakan 40 data <i>testing</i> menunjukkan nilai

Penulis	Judul Penelitian	Metode Penelitian	Hasil
			akurasi yang baik yaitu sebesar 85%.
(Kurniawan, 2017)	Klasifikasi Berita Menggunakan Metode <i>Improved Naive bayes</i>	<i>Naive Bayes Classifier</i>	Pada penelitian ini menggunakan delapan kategori berita yaitu : ekonomi, entertainment, olahraga, teknologi, kesehatan, makanan, otomotif, dan travel dengan hasil nilai precision 0.96, <i>Recall</i> terbaik pada data <i>training</i> dengan nilai 0.78 dan f-measure sebesar 0.86. Jumlah data <i>training</i> mempengaruhi performa, karena semakin banyak data <i>training</i> yang digunakan maka semakin baik.
(Ramadhani, 2015)	Analisis Sentimen Menggunakan Metode <i>Naive bayes classifier</i> dengan Model Dokumen Bernoulli dan <i>Support Vector Machine</i>	Metode <i>Naive Bayes Classifier</i> dengan Model Dokumen Bernoulli dan <i>Support Vector Machine</i>	Didapatkan bahwa akurasi yang paling tinggi yaitu pada metode NBC (<i>Naive Bayes Classifier</i>) sebesar 62,60%. Jumlah dari data <i>training</i> 5412 dan data <i>testing</i> 2701 yang telah diklasifikasi, didapatkan hasil opini untuk kelas positif sebesar 44,50% dan opini untuk kelas negatif sebesar 55,50%.
(Suryono, Utami, & Lufthi, 2018)	Klasifikasi Sentimen pada <i>Twitter</i> dengan <i>Naive Bayes Classifier</i>	<i>Naive Bayes Classifier</i>	Didapatkan hasil uji coba sebanyak 3 kali, nilai akurasi yang pertama sebesar 64,95%, yang kedua sebesar 66,36% dan yang ketiga sebesar 66,79%. Dari proses klasifikasi didapatkan opini positif sebesar 28%, opini negatif sebesar 20%, dan opini netral sebesar 52%. Hasil yang diperoleh opini netral yang dikaitkan dengan

Penulis	Judul Penelitian	Metode Penelitian	Hasil
			topik Presiden Joko Widodo.
(Falahah & Nur, 2015)	Pengembangan Aplikasi <i>Sentiment Analysis</i> Menggunakan Metode <i>Naive bayes</i>	<i>Naive Bayes Classifier</i>	Hasil dari analisis tersebut yaitu menghasilkan nilai akurasi sebesar 73% dengan jumlah data <i>training</i> 100 negatif, 100 positif, dan 100 netral.
(Hamzah, 2012)	Klasifikasi Teks dengan <i>Naive bayes classifier</i> (NBC) untuk Pengelompokkan Teks Berita dan Abstrak Akademis	<i>Naive Bayes Classifier</i> (NBC)	Jumlah data yang digunakan sebanyak 1000 dokumen berita dan 450 dokumen abstrak. Hasil pengujian menunjukkan bahwa akurasi dokumen berita sebesar 91%, sedangkan pada dokumen abstrak nilai akurasinya sebesar 82%.
(Faishal Nuruz Zuhri, 2017)	Analisis Sentimen Masyarakat terhadap <i>Brand Smartfren</i> Menggunakan <i>Naive bayes classifier</i> di Forum Kaskus	<i>Naive Bayes Classifier</i>	Hasil dari penelitian ini yaitu dari 6338 data uji, 4049 berhasil terklasifikasi ke dalam sentimen positif dan 3233 data bersentimen negatif. Dengan nilai akurasi sebesar 98,40%