

BAB II

LANDASAN TEORI

2.1 *Machine Learning*

Machine Learning adalah ilmu (seni) pemrograman komputer, sehingga mereka bisa belajar dari data (Goleman dkk., 2019). Pembelajaran mesin merupakan sistem yang dapat belajar dengan sendirinya untuk memutuskan sesuatu tanpa harus berulang kali dilakukan pemrograman oleh manusia, sehingga komputer menjadi lebih pintar dan dapat belajar dari data yang dimilikinya.

Menurut (Retnoningsih dan Pramudita, 2020), *Machine learning* bekerja apabila tersedia data sebagai input untuk dilakukan analisis terhadap kumpulan data besar sehingga menemukan pola tertentu. Data adalah bahan input yang digunakan untuk melakukan pembelajaran (*training*) sehingga mesin dapat menghasilkan analisis yang benar. Pada *machine learning* terdapat data *training* dan data *testing*, data *training* digunakan melatih algoritma didalam *machine learning* sedangkan data *testing* digunakan untuk mengetahui performa algoritma dalam *machine learning*. Algoritma ini telah dilatih, yaitu saat menemukan data baru yang belum pernah diberikan dalam data pelatihan.

Machine learning menggunakan teknologi untuk memproses data besar dengan cara yang cerdas untuk menghasilkan hasil yang benar. Menurut teknik pembelajarannya, jenis-jenis pembelajaran mesin dibedakan menjadi *supervised learning*, *unsupervised learning*, *semi-supervised learning* dan *reinforcement learning*. *Supervised learning* adalah teknik *machine learning* yang menggunakan

kumpulan data berlabel (data training) untuk melakukan pembelajaran mesin, sehingga mesin dapat menggunakan fitur-fiturnya untuk mengidentifikasi label input untuk prediksi dan klasifikasi lebih lanjut. Sedangkan *unsupervised learning* merupakan teknik yang dilakukan dengan menarik kesimpulan berdasarkan input data labeled response (Buslim, 2019).

Model klasifikasi *machine learning* dibagi menjadi dua yaitu a) Melatih data lalu membedakannya kedalam class; b) Melakukan evaluasi terhadap objek yang tidak dikenal (Fahrizal dkk., 2020). Klasifikasi merupakan bagian *unsupervised learning*, yang memerlukan persiapan data yang baik serta pemilihan bahasa pemrograman yang mampu mendukung *machine learning*.

2.2 Klasifikasi

Klasifikasi adalah proses pengelompokan yang artinya mengumpulkan benda atau entitas yang sama dan memisahkan objek atau entitas yang tidak sama. Menurut (Sumarlin, 2015) Klasifikasi merupakan proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui.

Proses klasifikasi didasarkan pada empat komponen (Gorunescu, 2011).

1. *Kelas*

Variabel dependen yang berupa kategorikal yang merepresentasikan 'label' yang terdapat pada objek.

Contohnya: resiko penyakit jantung, resiko kredit, customer loyalty, jenis gempa.

2. *Predictor*

Variabel independen yang direpresentasikan oleh karakteristik (atribut) data.

Contohnya: merokok, minum alkohol, tekanan darah, tabungan, aset, gaji.

3. *Training dataset*

Satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.

4. *Testing dataset*

Berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi.

2.3 Adaptive Boosting

Boosting merupakan konsep *machine learning* dengan mengkombinasikan beberapa algoritma *classifier* yang lemah untuk membentuk suatu *classifier* yang kuat. Sedangkan *Adaptive boosting* merupakan algoritma *boosting* yang mampu menyesuaikan secara adaptif nilai error yang dihasilkan oleh *classifier* lemah untuk dijadikan acuan pada proses pelatihan *classifier* berikutnya. *AdaBoost* digunakan karena mudah diimplementasikan (Prasetio dan Susanti, 2019) dan fleksibel maka dapat dikombinasikan dengan berbagai algoritma.

Adapun tahapan-tahapan *Adaptive Boosting* yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Membagi data menjadi data latih dan data uji kedalam beberapa scenario.
2. Menginisialisasi atau mengaplikasikan bobot yang sama untuk setiap data latih yaitu $w_n = 1/N$ dengan N merupakan jumlah data latih.
3. Menentukan jumlah iterasi maksimal.

4. Menentukan prediksi kelas data uji dan menghitung ketepatan akurasi

2.4 Algoritma *K-Nearest Neighbor*

Menurut (Arifin, 2019) Algoritma *K-Nearest Neighbor* diartikan sebagai suatu metode yang dapat melakukan klasifikasi data terdekat berdasarkan tetangga atau data sebelumnya sebagai sample untuk menemukan hasil akhir. Algoritma ini termasuk kedalam kelompok algoritma instance-based learning (Diwahana dkk., 2019). K-NN dibagi menjadi dua proses, yaitu *training* dan klasifikasi atau *testing*. Pada proses *training*, K-NN melakukan penyimpanan vektor-vektor fitur serta klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data yang akan diuji coba (yang klasifikasinya tidak diketahui). Adapun langkah-langkah untuk menghitung metode Algoritma *K-Nearest Neighbor*:

- a. Menentukan Parameter K atau Jumlah tetangga paling dekat.
- b. Menghitung jarak *Euclid* atau *query instance* pada masing-masing objek terhadap data sampel yang diberikan.
- c. Mengurutkan objek-objek ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
- d. Mengumpulkan kategori Y atau Klasifikasi *K-Nearest Neighbor*
- e. Dengan menggunakan kategori *Nearest Neighbor* yang paling banyak maka dapat dilakukan prediksi nilai *query instance* yang telah dihitung.

Perhitungan jarak ketetanggaan menggunakan algoritma eucliden seperti yang ditunjukkan pada persamaan 1.

$$euclidean = \sqrt{((a_1 - b_1)^2 + \dots + (a_n - b_n)^2)} \quad (1)$$

Dimana $a = a_1, a_2, \dots, a_n$, dan $b = b_1, b_2, \dots, b_n$ mewakili n nilai atribut dari dua record. Sebuah titik akan dilakukan prediksi dari jenisnya berdasarkan klasifikasi terbanyak dari *neighbor* di sekitarnya,

2.5 Confusion Matrix

Confusion matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. *Confusion matrix* adalah media yang berguna untuk menganalisis seberapa baik *classifier* dapat mengenali tupel dari kelas yang berbeda (Sumiah and Mirantika, 2020).

Evaluasi menggunakan *confusion matrix* akan menghasilkan nilai *accuracy*, *precision*, dan *recall*. *Confusion matrix* merupakan tabel matrix yang terdiri dari dua kelas, yaitu kelas yang satu dianggap sebagai positif dan yang lainnya negatif.

Tabel 2. 1 Tabel Matriks

Klasifikasi yang benar	Diklasifikasikan sebagai	
	+	-
+	<i>True Positive</i>	<i>False Negative</i>
-	<i>False Positive</i>	<i>True Negative</i>

Keterangan :

1. True positive : jumlah record positif yang diklasifikasikan sebagai positif,
2. False positive : jumlah record negatif yang diklasifikasikan sebagai positif,
3. False negative : jumlah record positif yang diklasifikasikan sebagai negatif,

4. True negative : jumlah record negatif yang diklasifikasikan sebagai negative.

2.6 Feature Encoding

Dalam *Machine Learning*, terdapat hal yang harus diperhatikan yaitu bahwa sebagian algoritma bekerja lebih maksimal dengan data bertipe numerik untuk mencapai hasil yang lebih baik. Perkembangan data secara pesat membuat tipe data tidak hanya berupa numerik, namun juga berupa teks atau label pada suatu dataset. Oleh karena itu, dalam penelitian perlu dilakukan mengubah data menjadi numerik untuk membuat sebuah model. Salah satu teknik yang dapat dilakukan adalah feature encoding yaitu teknik mengkonversi variabel teks atau kategorik menjadi nilai numerik. Terdapat dua cara yang paling umum dilakukan dengan menggunakan *Label Encoder* atau *One Hot Encoder* sebagai berikut:

1. *Label Encoding*: metode ini mengacu pada konversi label ke dalam bentuk angka sehingga dapat mengubahnya menjadi bentuk yang dapat dibaca oleh mesin, langkah ini merupakan *pre-processing* data yang penting untuk dataset terstruktur dalam *supervised learning*.
2. *One Hot Encoding*: metode ini akan mengkonversi setiap labelnya menjadi 0 dan 1, dan setiap label akan dipecah menjadi beberapa kolom sesuai dengan banyak label.

2.8 State of The Art Bidang Penelitian

Tabel 2.2. Menunjukkan perbandingan penelitian sebelumnya yang berhubungan dengan focus penelitian komparasi algoritma klasifikasi. Terdapat

beberapa kesamaan serta perbedaan dari masing-masing penelitian. Hal ini dapat dilihat dari penggunaan metode serta algoritmanya.

Tabel 2. 2 *State of The Art* Penelitian Terkait

No	Nama Pengarang	Tahun	Judul	Isi Ringkasan	Hasil
1.	Elly Firasari, Nurul Khasanah, Umi Khultsum, Desiana Nur Kholifah, Rachman Komarudin, dan Wiwiek Widyastuty	2020	Comparation of K-Nearest Neighbor (K-NN) and Naive Bayes Algorithm for the Classification of the Poor in Recipients of Social Assistance	Pemerintah memiliki beberapa program bantuan yang dilaksanakan diantaranya bantuan Beras Miskin, Program Keluarga Harapan (PKH), Bantuan Langsung Tunai (BLT). Proses klasifikasi penerima bansos masih dilakukan secara manual yang dinilai kurang akurat saat mendapatkan hasil penerima bansos. Kemudian, dua metode yang memberikan hasil perhitungan terbaik adalah Naïve Bayes dan K-NN, yang digunakan untuk menghitung penerima bantuan sosial di desa Somokerto.	Algoritma klasifikasi Naïve Bayes memiliki akurasi lebih tinggi dibandingkan algoritma K-NN dalam klasifikasi penduduk miskin dalam penerima bansos di Desa Somokerto. Algoritma Klasifikasi Naive Bayes memberikan hasil klasifikasi dengan akurasi 89,04% sedangkan algoritma K-NN menghasilkan klasifikasi dengan akurasi 87,67%. Perhitungan dua algoritma tersebut termasuk dalam kategori klasifikasi yang baik.

2.	Aah Sumiah, Nita Mirantika	2020	Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan	Pada penelitian ini membandingkan dua algoritma untuk mengetahui algoritma yang paling cocok digunakan dalam rekomendasi penentuan mahasiswa penerima beasiswa menggunakan algoritma K-Nearest Neighbor dan algoritma Naive Bayes. algoritma ini di pilih dikarenakan dua algoritma tersebut adalah algoritma yang populer digunakan dalam proses pengklasifikasian.	Perbandingan metode <i>K-Nearest Neighbor</i> dan <i>Naive Bayes</i> untuk rekomendasi penentuan mahasiswa penerima beasiswa ini menghasilkan bahwa algoritma <i>K-Nearest Neighbor</i> mempunyai akurasi yang lebih tinggi yaitu 100% dibandingkan algoritma <i>Naive Bayes</i> yang menghasilkan akurasi sebesar 99,89%.
3.	Devi Yunita	2017	Perbandingan Algoritma K-Nearest Neighbor Dan Decision Tree Untuk Penentuan Risiko Kredit Kepemilikan Mobil	Penelitian analisa kredit ini menggunakan perbandingan dari Algoritma K-nearest neighbor (K-NN) dimana penelitian ini menggunakan metode dengan mencari kedekatan kriteria kasus baru dengan kriteria kasus lama	Perbandingan dua algoritma yaitu algoritma <i>K-Nearest Neighbor</i> dan <i>Decision Tree</i> untuk penentuan risiko kredit kepemilikan mobil ini menggunakan aplikasi Rapid Miner dan hasilnya algoritma <i>K-</i>

				yang didasarkan pada kriteria kasus yang mendekati, dan menggunakan Metode Decision tree yaitu metode yang ada pada teknik klasifikasi dalam data mining.	<i>Nearest Neighbor</i> memiliki nilai akurasi yang lebih baik yaitu dengan akurasi 98.18%.
4.	Ade Irma Prianti, Rukun Santoso, Arief Rachman Hakim	2020	Perbandingan Metode K-Nearest Neighbor Dan Adaptive Boosting Pada Kasus Klasifikasi Multi Kelas	Perbandingan metode <i>K-Nearest Neighbor</i> dan <i>Adaptive Boosting</i> dengan menggunakan <i>base learner</i> -nya yaitu <i>Classification and Regression Trees</i> (CART) pada kasus klasifikasi multi kelas.	Hasil dari perbandingan metode <i>K-Nearest Neighbor</i> dan <i>Adaptive Boosting</i> tersebut mendapatkan model terbaik dengan akurasi tertinggi yaitu pada metode <i>Adaboost</i> dengan banyak nilai <i>accuracy</i> , <i>precision</i> serta <i>recall</i> yang lebih besar dibandingkan dengan nilai pada model algoritma <i>K-Nearest Neighbor</i> .
5.	Riyan Latifahul Hasanah, Muhamad Hasan, Witriana	2019	Klasifikasi Penerima Dana Bantuan Desa Menggunakan Metode	Status keluarga miskin sebagai penerima bantuan sangat penting dilakukan agar bantuan	Penelitian dilakukan melalui tahap perhitungan manual untuk mengetahui kelayakan atau

	Endah Pangesti, Fanny Fatma Wati, dan Windu Gata		KNN (K-Nearest Neighbor)	penanggulangan kemiskinan dari pemerintah dapat dialokasikan dengan tepat, maka dari permasalahan tersebut dilakukanlah pengujian K-Fold Cross Validation pada algoritma K-Nearest Neighbors untuk memprediksi penerimaan dana bantuan desa.	ketidaklayakan data baru, dan aplikasi Rapidminer digunakan untuk menguji keakuratan kumpulan data dengan berbagai nilai K. Hasilnya ialah dengan K=15 dan K=30 data yang baru (D160) memiliki kategori “Tidak Layak” dengan akurasi sebesar 100%. Kemudian dengan nilai K=45, K=60 dan K=75 data yang baru (D160) memiliki kategori “Layak” dengan akurasi sebesar 81,25%.
6.	Rizki Tri Prasetio, Sari Susanti	2019	Prediksi Harapan Hidup Pasien Kanker Paru Pasca Operasi Bedah Toraks Menggunakan Boosted k-Nearest Neighbor	Kanker paru-paru menjadi salah satu penyakit penyebab kematian terbanyak di Indonesia. Faktor penyebabnya oleh asap rokok. Operasi toraks adalah salah satu solusi utama untuk kanker paru-	Hasil menunjukkan bahwa metode yang digunakan menghasilkan tingkat akurasi prediksi harapan hidup sebesar 85.11% yang menggunakan validasi 10 fold cross validation dengan

				<p>paru. Namun, ada banyak risiko dan komplikasi setelah operasi toraks yang berujung pada kematian. Maka akan dilakukan prediksi harapan hidup pasien. Prediksi dilakukan dengan menganalisis kondisi pasien sebelum dan sesudah operasi. Adaptive Boost digunakan sebagai optimasi level algoritma pada algoritma k-nearest neighbor.</p>	<p>parameter nilai k pada algoritma k-nearest neighbor bernilai 5</p>
7.	Aah Sumiah, dan Nita Mirantika	2020	Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan	<p>Membandingkan algoritma K-Nearest Neighbor dan algoritma Naive Bayes untuk rekomendasi penerima beasiswa pada universitas kuningan. Alasan dipilihnya metode ini adalah karena kedua algoritma ini biasa digunakan dalam proses klasifikasi data. Hasilnya data di implementasikan menjadi sistem</p>	<p>Hasil dari algoritma K-Nearest Neighbor diperoleh akurasi sebesar 100% dan Algoritma Naive Bayes diperoleh akurasi sebesar 99,89%. Hal ini menunjukkan bahwa pada data penerimaan beasiswa ini tingkat akurasi K-nearest Neighbor lebih tinggi dibandingkan dengan</p>

				informasi yang menggunakan visual basic.net dan sql server.	Naive Bayes sehingga lebih cocok digunakan algoritma K-nearest neighbor untuk rekomendasi penerima beasiswa. Banyaknya data latih akan mempengaruhi keakuratan data.
8.	Hilda Nur Zerlinda, Isnandar Slamet, dan Etik Zukhronah	2019	Klasifikasi Calon Penerima Bidikmisi Dengan Menggunakan Algoritma K-Nearest Neighbor	Pelamar Bidikmisi UNS selalu bertambah setiap tahunnya, namun kuota peserta yang diterima tidak sebanding dengan jumlah pendaftar. Maka dilakukanlah klasifikasi dengan algoritma K-Nearest Neighbor untuk mengklasifikasikan calon penerima UNS Bidikmisi. Dalam penelitian ini memiliki beberapa variabel bebas seperti pendapatan orang tua, pekerjaan ibu orang tua, jumlah tanggungan, luas tanah, sumber air dan kepemilikan	Penggunaan algoritma KNN untuk membandingkan hasil prediksi klasifikasi dengan data observasi bidikmisi. Data yang digunakan sebanyak 1539 data latih dan 500 data uji. Jadi dapat disimpulkan bahwa hasil perhitungan yang diterapkan menghasilkan nilai akurasi sebesar 84,4% Untuk $k = 5$.

				rumah. Tahapan pelaksanaan penelitian ini adalah menormalisasi data, membagi data menjadi data latih dan data uji, serta menentukan nilai k.	
9.	Dita Noviana, Yuliana Susanti, Irwan Susanto	2019	Analisis Rekomendasi Penerima Beasiswa Menggunakan Algoritma K-Nearest Neighbor (K-Nn) Dan Algoritma C4.5	Dengan semakin banyaknya mahasiswa yang mengajukan beasiswa, metode klasifikasi diperlukan untuk membantu menentukan siapa yang berhak mendapatkan beasiswa PPA di Universitas Sebelas Maret. Metode yang digunakan dalam analisis ini adalah K-nearest neighbor dan C4.5. Hasil klasifikasi akan digunakan untuk pengambilan keputusan rekomendasi penerima beasiswa.	Hasil dari algoritma k-NN dan C4.5 dengan akurasi masing-masing sebesar 90.7% dan 88.3%. Hasil penelitian menunjukkan bahwa variabel yang berpengaruh adalah IPK, prestasi, penghasilan, jumlah orang tua bekerja dan jumlah tanggungan. Variabel yang paling berpengaruh adalah variabel IPK.

10.	Erlangga Dwi Kurniawan, dan Mufti	2018	Implementasi Algoritma K-Nearest Neighbor Dengan Metode Klasifikasi Dan Pengukuran Jarak Manhattan Distance Untuk Prediksi Kelulusan Un Berdasarkan Hasil Nilai Tryout Berbasis Java Desktop Pada Sma Harapan Jaya 2	Dalam memprediksi kelulusan ujian nasional, sekolah masih melakukannya secara manual yaitu menggunakan Microsoft Excel. Berdasarkan hal tersebut, melalui algoritma K-Nearest dan metode jarak "Manhattan Distance" dapat diprediksi bahwa siswa yang lulus ujian nasional akan terbagi menjadi dua kelas yaitu Lulus dan Tidak Lulus. Atribut tersebut akan mempengaruhi kinerja akademik siswa.	Nilai akurasi pengujian algoritma K-Nearest Neighbor dengan pengukur jarak Manhattan Distance berdasarkan jumlah uji yang dilakukan sebanyak 3 kali dengan jumlah 201 jurusan IPA dan 210 jurusan IPS selama 3 angkatan didapatkan nilai $K = 5$ untuk jurusan IPA dan nilai $K = 3$ untuk jurusan IPS dengan nilai rata-rata tingkat akurasi tertinggi 89.08% di tahun 2014/2015 untuk jurusan IPA dan nilai rata-rata tingkat akurasi tertinggi 97.30% ditahun 2015/2016 untuk jurusan IPS.
11.	Sumarlin	2015	Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung	Dengan perkemebangan bidang akademik khususnya perguruan tinggi, beasiswa adalah masalah	Hasil performa algoritma knearest neighbor menggunakan metode cross validation,

			<p>Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM</p>	<p>yang menarik untuk diteliti. Penelitian ini membahas tentang klasifikasi beasiswa PPA dan BBM berdasarkan variabel yang telah ditentukan dengan menerapkan algoritma k-nearest neighbour. Proses pemilihan penerima beasiswa untuk meningkatkan prestasi akademik dan bantuan belajar mahasiswa memerlukan sistem pendukung keputusan (SPK) untuk membantu memberikan alternatif solusi.</p>	<p>Confusion Matrix dan kurva Receiver Operating Characteristic (ROC), akurasi yang dihasilkan untuk beasiswa PPA mencapai 88,33% dengan nilai Area Under Curva 0,925 dari 227 record dataset, sedangkan akurasi yang dihasilkan untuk beasiswa BBM mencapai 90% dengan nilai AUC 0,937 dari 183 record dataset, akurasi yang dihasilkan untuk gabungan beasiswa PPA dan BBM mencapai 85,56% dan nilai AUC sebesar 0,958. dikarenakan nilai AUC ada pada rentang 0,9 sampai 1,0 maka masuk kedalam kategori sangat baik.</p>
--	--	--	--	---	--

12.	Moh. Syaiful Anam	2021	Sistem Pendukung Keputusan Sosial Menggunakan Metode Naive Bayes	Dalam masa pandemi covid -19, pandemi ini menyebar ke seluruh sendi kehidupan dan salah satu yang paling menjadi perhatian adalah dibidang sosial ekonomi. Banyak terdapat bantuan Sosial (Bansos) yang disalurkan baik oleh pemerintah ataupun pihak swasta lain. Penelitian ini bertujuan untuk membuat sistem pendukung keputusan bantuan sosial menggunakan metode Naive Bayes, selanjutnya melakukan Analisa menggunakan tabel Confusion Matrix.	Berdasarkan hasil pengujian confusion matrix dengan teknik split validasi, penggunaan metode klasifikasi naïve bayes terhadap dataset yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori Good. Sementara nilai Precision sebesar 92% dan Recall sebesar 86%. Berdasarkan hal tersebut dapat dinyatakan bahwa sistem klasifikasi yang dibangun dapat gunakan sebagai bahan masukan bagi pengambil keputusan.
-----	-------------------	------	--	---	--

Pada *state of the art* tersebut belum ada penelitian yang meneliti membandingkan dua algoritma klasifikasi terbaik yang dihasilkan pada penelitian sebelumnya yaitu algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* dalam menentukan penerima

bantuan sosial. Algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* masing-masing memiliki kekurangan dan kelebihan. Oleh karena itu, pada penelitian ini akan melakukan perbandingan dari kedua algoritma tersebut untuk memperoleh hasil algoritma yang paling maksimal dalam menentukan penerima bantuan sosial.

2.9 Matriks Penelitian

Tabel 2. 3 Tabel Matriks Penelitian

No.	Penulis/Tahun	Judul	Ruang Lingkup									
			Algoritma/Metode				Tujuan		Objek			
			Naïve Bayes	KNN	Decision Tree	Ada boost	Klasifikasi	Prediksi	Ban sos	Bea siswa	Penyakit	Lainnya
1.	(Anam, 2021)	Sistem Pendukung Keputusan Bantuan Sosial Dengan Menggunakan Metode Naive Bayes	√	-	-	-	-	-	√	-	-	-
2.	(Firasari dkk., 2020)	Comparison of K-Nearest Neighbor (K-NN) and Naive Bayes Algorithm for the Classification of the Poor in Recipients of Social Assistance	√	√	-	-	√	-	√	-	-	-
3.	(Sumiah dan Mirantika, 2020)	Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk	√	√	-	-	-	√	-	√	-	-

		Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan											
4.	(Yunita, 2017)	Perbandingan Algoritma K-Nearest Neighbor Dan Decision Tree Untuk Penentuan Risiko Kredit Kepemilikan Mobil	-	√	√	-	-	√	-	-	-	-	√
5.	(Prianti dkk., 2020)	Perbandingan Metode K-Nearest Neighbor Dan Adaptive Boosting Pada Kasus Klasifikasi Multi Kelas	-	√	-	√	√	-	-	-	-	-	√
6.	(Yusi, 2021)	Analisa Perbandingan Kinerja Algoritma <i>K-Nearest Neighbor</i> Dan <i>Adaptive Boosting</i> Pada Prediksi Penerima Bantuan Sosial	-	√	-	√	-	√	√	-	-	-	-

2.10 Relevansi Penelitian

Tabel 2. 4 Tabel Relevansi Penelitian

Peneliti	(Prianti dkk., 2020)	(Yusi Yustikasari, 2021)
Judul	Perbandingan Metode K-Nearest Neighbor Dan Adaptive Boosting Pada Kasus Klasifikasi Multi Kelas	Analisa Perbandingan Kinerja Algoritma <i>K-Nearest Neighbor</i> Dan <i>Adaptive Boosting</i> Pada Prediksi Penerima Bantuan Sosial Pangan Non Tunai
Masalah Penelitian	Melakukan perbandingan antara metode K-Nearest Neighbor dan Adaptive Boosting untuk mengetahui metode yang lebih baik dalam memprediksi kinerja perusahaan di Indonesia yang merupakan kasus klasifikasi multi kelas	Membandingkan kinerja dua algoritma klasifikasi terbaik yang dihasilkan pada penelitian sebelumnya yaitu algoritma <i>K-Nearest Neighbor</i> dan <i>Adaptive Boosting</i> dalam menentukan penerima bantuan sosial pangan non tunai.
Objek Penelitian	Klasifikasi Multi Kelas	Klasifikasi Binary Class
Algoritma / Metode	<i>K-Nearest Neighbor</i> dan <i>Adaptive Boosting</i> dengan <i>base learning Adaboost</i> menggunakan algoritma CART	<i>K-Nearest Neighbor</i> dan <i>Adaptive Boosting</i> dengan <i>base learning Adaboost</i> menggunakan algoritma <i>Decision Tree</i>
Implementasi	Klasifikasi dilakukan dengan <i>Tools RapidMiner</i>	Prediksi dilakukan dengan model <i>Machine Learning</i> bahasa pemrograman <i>Python</i>