



**PERBANDINGAN METODE DATA MINING
K-NEAREST NEIGHBORS DAN NAÏVE BAYES UNTUK PREDIKSI
KESIAPAN ANAK DALAM MENYESUAIKAN DIRI DI JENJANG
SEKOLAH DASAR**

TUGAS AKHIR

AGUSTINUS BONA VENTURA
41516120017

UNIVERSITAS
MERCU BUANA
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021**



**PERBANDINGAN METODE DATA MINING
K-NEAREST NEIGHBORS DAN NAÏVE BAYES UNTUK PREDIKSI
KESIAPAN ANAK DALAM MENYESUAIKAN DIRI DI JENJANG
SEKOLAH DASAR**

TUGAS AKHIR

AGUSTINUS BONA VENTURA
41516120017

UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41516120017

Nama : Agustinus Bona Ventura

Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 18 Juli 2021



Agustinus Bona Ventura



SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Agustinus Bona Ventura
NIM : 41516120017
Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Juli 2021

UNIVERSITAS
MERCU BUANA



Agustinus Bona Ventura

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Agustinus Bona Ventura
 NIM : 41516120017
 Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi ✓	
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal	: JUITA JURNAL INFORMATIKA	
	ISSN	: 2579-8901	
	Link Jurnal	:	
	Link File Jurnal Jika Sudah di Publish	:	

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui

Jakarta, 18 Juli 2021


 (Dr. Leonard Goeirmento)
 Dosen Pembimbing



Agustinus Bona Ventura

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120017
Nama : Agustinus Bona Ventura
Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 Februari 2021



(Desi Ramayanti, S.Kom., MT)
Ketua Penguji

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120017
Nama : Agustinus Bona Ventura
Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 Februari 2021



(Anis Cherid, SE, MTI)
Anggota Penguji 1

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120017
Nama : Agustinus Bona Ventura
Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 Februari 2021



(Afiyati, S.Si, MT)
Anggota Penguji 2

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41516120017
Nama : Agustinus Bona Ventura

Judul Tugas Akhir : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 19 Februari 2021

Menyetujui,


(Dr. Leonard Goeirmanto)
Dosen Pembimbing

Mengetahui,

UNIVERSITAS
MERCU BUANA


(Diky Firdaus, S.Kom, MM)

Koord. Tugas Akhir Teknik Informatika


(Desi Ramadanti, S.Kom, MT)

Ka. Prodi Teknik Informatika

ABSTRAK

Nama : Agustinus Bona Ventura
NIM : 41516120017
Pembimbing TA : Leonard Goeirmanto
Judul : Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar

Dalam peraturan Permendikbud nomor 17 tahun 2017 pasal 5 ayat (1) dan (2) yang menguraikan persyaratan batas usia calon siswa didik kelas 1 (satu) SD adalah 7 tahun dan jika calon peserta didik berusia dibawah 6 tahun, maka calon siswa tersebut membutuhkan rekomendasi tertulis dari Psikolog yang berisi keterangan calon siswa tersebut dan belajar di SD. Kenyataannya, banyak orang tua tidak sabar untuk dan membiarkan anak-anak mereka memasuki sekolah dasar walaupun usianya tidak mencukupi, maka dari itu kesiapan sekolah dapat diukur melalui matangnya aspek perkembangan yang secara langsung dan membantu anak dalam proses pembelajaran dan penyesuaian dirinya dilingkungan sekolah. Adapun tujuan dari penelitian ini untuk memprediksi kesiapan anak masuk sekolah dasar dengan menggunakan metode *naïve bayes* dan *K-Nearest Neighbors*, atribut yang digunakan terdiri dari Perkembangan Nilai Agama dan Moral yang terdiri dari 11 katogori, Perkembangan Fisik - Motorik yang terdiri dari 22 kategori, Perkembangan Kognitif yang terdiri dari 20 kategori, Perkembangan Sosial - Emosional yang terdiri dari 16 kategori, Perkembangan Bahasa yang terdiri dari 14 kategori, Perkembangan seni yang terdiri dari 16 kategori. Untuk melakukan proses data *mining* tersebut diperlukan *tools* pembantu yaitu *RapidMiner*. adapun pengujian sistem terdiri dari pengujian pengaruh nilai K dan pengujian pengaruh nilai K-Fold. Hasil dari pengujian pengaruh nilai K-Fold menghasilkan akurasi *naïve bayes* optimum senilai 54.63% +/- 4.24% (micro average : 54.63%) pada nilai K-Fold 3, sedangkan pengujian nilai akurasi K-Nearest Neighbors optimum senilai 62.96% +/- 2.87% (micro average : 62.96%) dengan nilai K=10 pada nilai K-Fold 6. Dari hasil tersebut, disimpulkan bahwa metode K-Nearest Neighbors dapat digunakan dengan baik untuk memprediksi kesiapan anak masuk sekolah dasar.

Kata Kunci: Kesiapan Anak Sekolah Dasar, Data Mining, Naïve Bayes, K-Nearest Neighbors, Rapidminer.

ABSTRACT

Name : Agustinus Bona Ventura
Student Number : 41516120017
Counsellor : Leonard Goeirmento
Title : The Comparison of Data Mining K-Nearest Neighbors and Naïve Bayes Methodology to Predict Children's Readiness in Adjusting Themselves in Elementary School Environment

According to the revocation of Education and Culture Minister Regulation number 17/2015 5:1-2, the minimum age of grade 1 of elementary student is 7 years old and if a student candidate is under 6 years old, the student candidate is required to have written recommendation from a psychologist to enroll in elementary school. In fact, many parents insist to early enroll their underage-children in elementary school. In response to this issue, the readiness of a school is shown by providing children development care to assist the children in learning and adjusting themselves in school environment. The objective of this research is to predict student candidates' readiness to enroll elementary school following naïve bayes dan K-Nearest Neighbors methodology and considering 11 categories of religious and moral aspect, 22 categories of physical – motoric aspect, 20 categories of cognitive aspect, 16 categories of social-emotional development, 14 categories of language development, and 16 categories of art aspect. To process the mining data, RapidMiner is needed. As for the system testing, it consists of testing the effect of the K value and testing the effect of the K-Fold value. The results of testing the effect of the K-Fold value produce an optimum accuracy of naïve Bayes 54.63% +/- 4.24% (micro average : 54.63%) at the K-Fold 3 value while the optimum accuracy of K-Nearest Neighbors is 62.96% +/- 2.87% (micro average : 62.96%) with K=10 at the K-Fold 6. From the result, can be concluded that K-Nearest Neighbors methodology can be used to predict the readiness of children to enroll elementary school.

Key word: children's readiness to enroll elementary school, Mining data, Naïve Bayes, K-Nearest Neighbors, Rapidminer

KATA PENGANTAR

Puji syukur kita panjatkan kehadiran Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi dengan judul *“Perbandingan metode Data Mining Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri di Jenjang Sekolah Dasar Dengan Menggunakan Metode Naïve Bayes”*. Tujuan penulisan skripsi ini untuk memenuhi syarat memperoleh gelar Sarjana Komputer (S.Kom.) bagi mahasiswa program S-1 di program studi Ilmu Komputer Jurusan Teknik Informatika Universitas Mercu Buana. Penulis menyadari bahwa tanpa bantuan dan bimbingan Leonard Goeirmanto. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Orang tua yang selalu memberikan dukungan dan doa sehingga penulis dapat menyelesaikan penulisan jurnal dan tugas akhir beserta laporannya dengan lancar dan tepat waktu.
2. Bapak Leonard Goeirmanto selaku dosen pembimbing tugas akhir yang telah meluangkan waktu dan pikirannya untuk membimbing penulis sehingga dapat menyelesaikan tugas akhir dengan lancar.
3. Bapak Achmad Kodar, Drs. MT, selaku dosen pembimbing akademik yang telah membimbing saya sejak awal semester yang selalu memberi dukungan dan motivasi agar lulus tepat waktu.
4. Bapak/Ibu dosen Fakultas Ilmu Komputer Universitas Mercu Buana Meruya yang telah memberikan ilmu dan membimbing agar bisa menjadi mahasiswa yang berguna bagi orang lain.
5. Kepala Sekolah, Guru dan Psikolog sekolah TB – TK Santa Ursula Jakarta.
6. Semua teman-teman saya yang telah memberikan semangat dan dukungan serta membantu saya dalam menyusun laporan tugas akhir ini.

Akhir kata, penulis mengucapkan terimakasih kepada semua pihak yang telah membantu dan penulis berharap semoga skripsi ini dapat bermanfaat bagi kita semua dan menjadi bahan masukan dalam dunia pendidikan.

Jakarta, 8 Februari 2021
Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR	iv
LEMBAR PERSETUJUAN PENGUJI.....	v
LEMBAR PENGESAHAN.....	viii
ABSTRAK.....	ix
ABSTRACT	x
KATA PENGANTAR	xi
DAFTAR ISI.....	xii
NASKAH JURNAL.....	1
KERTAS KERJA	22
BAB 1. LITERATUR REVIEW	23
BAB 2. ANALISIS DAN PERANCANGAN	25
BAB 3. SOURCE CODE	34
BAB 4. DATASET	40
BAB 5. TAHAPAN EKSPERIMEN	43
BAB 6. HASIL SEMUA EKSPERIMEN	45
DAFTAR PUSTAKA	79
LAMPIRAN DOKUMEN HAKI	80
LAMPIRAN KORESPONDENSI	81

NASKAH JURNAL

Perbandingan Metode Data Mining K-Nearest Neighbors dan Naïve Bayes Untuk Prediksi Kesiapan Anak Dalam Menyesuaikan Diri Di Jenjang Sekolah Dasar

Agustinus Bona Ventura¹, Leonard Goeirmanto²

Program Teknik Informatika Fakultas Ilmu Komputer Universitas Mercu Buana

41516120017@student.mercubuana.ac.id , tugas@dataku.web.id

Abstract

According to the revocation of Education and Culture Minister Regulation number 17/2015 5:1 -2, the minimum age of grade 1 of elementary student is 7 years old and if a student candidate is under 6 years old, the student candidate is required to have written recommendation from a psychologist to enroll in elementary school. In fact, many parents insist to early enroll their underage-children in elementary school. In response to this issue, the readiness of a school is shown by providing children development care to assist the children in learning and adjusting themselves in school environment. The objective of this research is to predict student candidates' readiness to enroll elementary school following naïve bayes dan K-Nearest Neighbors methodology and considering 11 categories of religious and moral aspect, 22 categories of physical – motoric aspect, 20 categories of cognitive aspect, 16 categories of social-emotional development, 14 categories of language development, and 16 categories of art aspect. To process the mining data, RapidMiner is needed. As for the system testing, it consists of testing the effect of the K value and testing the effect of the K-Fold value. The results of testing the effect of the K-Fold value produce an optimum accuracy of naïve Bayes 54.63% +/- 4.24% (micro average : 54.63%) at the K-Fold 3 value while the optimum accuracy of K-Nearest Neighbors is 62.96% +/- 2.87% (micro average : 62.96%) with K=10 at the K-Fold 6. From the result, can be concluded that K-Nearest Neighbors methodology can be used to predict the readiness of children to enroll elementary school.

Key word: children's readiness to enroll elementary school, Mining data, Naïve Bayes, K-Nearest Neighbors, Rapidminer.

Abstrak

Dalam peraturan Permendikbud nomor 17 tahun 2017 pasal 5 ayat (1) dan (2) yang menguraikan persyaratan batas usia calon siswa didik kelas 1 (satu) SD adalah 7 tahun dan jika calon peserta didik berusia dibawah 6 tahun, maka calon siswa tersebut membutuhkan rekomendasi tertulis dari Psikolog yang berisi keterangan calon siswa tersebut dan belajar di SD. Kenyataannya, banyak orang tua tidak sabar untuk dan membiarkan anak-anak mereka memasuki sekolah dasar walaupun usianya tidak mencukupi, maka dari itu kesiapan sekolah dapat diukur melalui matangnya aspek perkembangan yang secara langsung dan membantu anak dalam proses pembelajaran dan penyesuaian dirinya di lingkungan sekolah. Adapun tujuan dari penelitian ini untuk memprediksi kesiapan anak masuk sekolah dasar dengan menggunakan metode naïve bayes dan K-Nearest Neighbors, atribut yang digunakan terdiri dari Perkembangan Nilai Agama dan Moral yang terdiri dari 11 katogori, Perkembangan Fisik - Motorik yang terdiri dari 22 kategori, Perkembangan Kognitif yang terdiri dari 20 kategori, Perkembangan Sosial - Emosional yang terdiri dari 16 kategori, Perkembangan Bahasa yang terdiri dari 14 kategori, Perkembangan seni yang terdiri dari 16 kategori. Untuk melakukan proses data mining tersebut diperlukan tools pembantu yaitu RapidMiner. adapun pengujian sistem terdiri dari pengujian pengaruh nilai K dan pengujian pengaruh nilai K-Fold. Hasil dari pengujian pengaruh nilai K-Fold menghasilkan akurasi naïve bayes optimum senilai 54.63% +/- 4.24% (micro average : 54.63%) pada nilai K-Fold 3, sedangkan pengujian nilai akurasi K-Nearest Neighbors optimum senilai 62.96% +/- 2.87% (micro average : 62.96%) dengan nilai K=10 pada nilai K-Fold 6. Dari hasil tersebut, disimpulkan bahwa metode K-Nearest Neighbors dapat digunakan dengan baik untuk memprediksi kesiapan anak masuk sekolah dasar.

Kata Kunci: Kesiapan Anak Sekolah Dasar, Data Mining, Naïve Bayes, K-Nearest Neighbors, Rapidminer.

I. PENDAHULUAN

I.I. Latar Belakang

Perkembangan kehidupan anak akan selalu melibatkan masa peralihan, salah satunya adalah masa peralihan dalam jenjang pendidikan. Pendidikan dijadikan sebagai alat yang strategis dalam pembentukan dan pengembangan nilai, sikap dan moral seseorang.

Anak-anak yang mengikuti pendidikan di Taman Kanak-Kanak mudah dilewati karena banyaknya perubahan yang akan dihadapi oleh anak, misalnya saja cara belajar yang lebih formal, materi pelajaran yang lebih kompleks, atau aturan-aturan baru yang harus mereka ikuti.

Kesiapan sekolah sangat diperlukan dan perlu dipersiapkan orangtua serta pengajar di PAUD/TK agar anak yang akan melanjutkan ke SD dapat beadaptasi dengan perubahan-perubahan. Hanya saja pada masa peralihan ini seringkali orang tua beranggapan bahwa usia serta Taman Kanak-kanak (TK) atau Pendidikan Anak Usia Dini (PAUD) akan menghadapi masa peralihan ke jenjang pendidikan formal yaitu Sekolah Dasar (SD). Masa peralihan ini merupakan fase yang penting dan kemampuan anak untuk membaca, menulis dan berhitung (Calistung) merupakan indikator utama dalam melihat kesiapan anak mereka untuk melanjutkan pendidikan SD.

Umumnya orang tua mendaftarkan anaknya karena mereka meyakini usia anak mereka telah siap. Kesiapan ini dilihat dari terselesaikannya pendidikan TK. Pertimbangan tersebut perlu dikaji ulang, karena pada kenyataannya terdapat peraturan Permendikbud nomor 17 tahun 2017 pasal 5 ayat (1) dan (2) yang menguraikan persyaratan batas usia calon siswa didik kelas 1 (satu) SD adalah 7 tahun dan jika calon siswa didik berusia di bawah 6 tahun, maka calon siswa tersebut membutuhkan rekomendasi tertulis dari Psikolog yang berisi keterangan calon siswa tersebut telah siap untuk sekolah dan belajar di SD.

I.II. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diungkapkan di atas, maka terdapat rumusan masalah sebagai berikut :

- 1) Apakah metode *Naïve Bayes* dan *K-Nearest Neighbors* dapat digunakan untuk mengklasifikasikan data prediksi kesiapan anak dalam menyesuaikan diri di jenjang sekolah dasar ?
- 2) Bagaimana mengimplementasikan metode *Naïve Bayes* dan *K-Nearest Neighbors* untuk memprediksi kesiapan anak dalam menyesuaikan diri di jenjang sekolah dasar ?

I.III. Tujuan Penelitian

Tujuan dari penelitian ini diharapkan untuk mengetahui kesiapan TK Santa Ursula Jakarta dalam mempersiapkan menuju ke jenjang Sekolah Dasar.

I.IV. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut :

- 1) Hanya memprediksi kesiapan anak untuk ke jenjang sekolah dasar (SD).
- 2) Menggunakan metode *Naïve Bayes* dan *K-Nearest Neighbors*.

I.V. Landasan Teori

Berikut ini dasar teori yang terdapat pada penelitian :

1. Penelitian Terkait

Penelitian yang berkaitan dengan memprediksi kesiapan anak dalam menyesuaikan diri di jenjang sekolah dasar adalah sebagai berikut:

- 1). (Intan 2019) Penelitian ini berjudul “SISTEM PREDIKSI PRE-TEST PENENTUAN KESIAPAN MEMASUKI PENDIDIKAN DASAR” Data yang menjadi acuan sebagai proses mining sebanyak 200 sampel data anak pada Biro Konsultasi Dwipayana Makasar dua tahun terakhir yang sudah menyelesaikan test kesiapan sekolah dinyatakan siap sekolah, ragu dan belum siap. Berdasarkan hasil yang diperoleh pada pembangunan system dan pengujian perangkat lunak maka dapat disimpulkan bahwa Aplikasi data mining ini dapat memprediksi kesiapan sekolah anak usia 5 tahun samapai tuju tahun, yang siap sekolah , ragu dan belum siap sekolah dengan menggunakan algoritma K-Nearest Neighbor sebesar 68,3%

- 2). (Widaningsih 2019) Penelitian ini berjudul “PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA DENGAN ALGORITMA C4.5, NAÏVE BAYES, KNN, DAN SVM”. Data yang menjadi acuan dalam penelitian ini adalah data mahasiswa Teknik Informatika pada tahun 2008 hingga 2013. Data nilai berasal dari transkrip nilai mahasiswa yang terdapat pada database bagian akademik Fakultas Teknik. Sedangkan data kelulusan berasal dari data wisuda mahasiswa setiap tahunnya. Nilai k diambil 10 fold sehingga dari data 466 data akan menjadi 10 subset data dengan ukuran sama yaitu sekitar 46,6 atau 47 data. Dari masing-masing 10 subset tersebut, 419 data menjadi data latih dan 47 data menjadi data uji. Hasil performasi pada setiap model yaitu accuracy, error dan ROC selanjutnya akan dibandingkan untuk mengetahui algoritma mana yang lebih baik. Dari hasil perbandingan algoritma Naïve Bayes memiliki nilai yang paling baik untuk semua kategori performasi dibandingkan dengan algoritma lainnya

yaitu Accuracy 76.79%, Error 23.17% dan AUC 0.850.

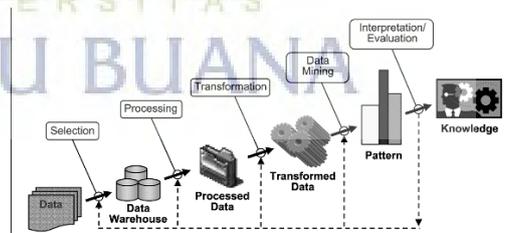
- 3). (Rahman, Hidayat, and Affif Supianto 2018) Penelitian ini berjudul “KOMPARASI METODE DATA MINING K-NEAREST NEIGHBOR DENGAN NAÏVE BAYES UNTUK KLASIFIKASI KUALITAS AIR BERSIH (STUDI KASUS PDAM TIRTA KENCANA KABUPATEN JOMBANG)” Data yang digunakan dalam penelitian merupakan data kualitas air bersih yang diperoleh dari kantor PDAM Tirta Kencana Kabupaten Jombang dalam jangka waktu tahun 2016 hingga 2017. Pada pengujian berdasarkan nilai atribut K didapatkan rata-rata nilai akurasi tertinggi pada K bernilai 3 yaitu 90.73%. Pada pengujian berdasarkan percentage split nilai rata-rata keseluruhan metode K-Nearest Neighbor adalah 82.45% dan metode Naive Bayes sebesar 72.52%. Sedangkan pada pengujian berdasarkan jumlah data training metode K-Nearest Neighbor memiliki rata-rata akurasi keseluruhan sebesar 83.32% dan metode Naive Bayes sebesar 70.91%. Berdasarkan hasil akurasi yang didapatkan dari seluruh pengujian metode K-Nearest Neighbor merupakan metode yang terbaik dengan total rata-rata akurasi 82.89%.
- 4). (Putri, Suparti, and Rahmawati 2014) Penelitian ini berjudul “PERBANDINGAN METODE KLASIFIKASI NAÏVE BAYES DAN K-NEAREST NEIGHBOR PADA ANALISIS DATA STATUS KERJA DI KABUPATEN DEMAK TAHUN 2012” Data yang digunakan dalam penelitian ini adalah data Status Kerja di Kabupaten Demak tahun 2012. Data bersumber dari Surver Angkatan Kerja Nasional (SAKERNAS) yang berjumlah 1375. Berdasarkan hasil perhitungan Press’s Q dapat dikatakan bahwa pengklasifikasian status kerja di Kabupaten Demak tahun 2012 dengan metode Naïve Bayes dan metode K-Nearest Neighbor sudah baik atau sudah akurat. Perhitungan APER dieproleh nilai laju error untuk metode Naïve Bayes yaitu sebesar 0.0591 dan metode K-Nearrest Neighbor sebesar 0.0394. Dari kedua nilai tersebut dapat dikatakan bahwa baik metode Naïve Bayes ataupun metode K-Nearst Neighbor memliki peluang yang kecil untuk melakukan kesalahan dalam pengklasifikasian.

2. Teori Pendukung
 - 1) DataMining

(Widaningsih 2019) Data mining merupakan salah satu teknik untuk menggali atau “menambang” pengetahuan dari sekumpulan besar data. Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya yang dapat dipahami dan bermanfaat bagi pemilik data (Larose, 2005). Terdapat beberapa teknik yang digunakan untuk data mining seperti yang diungkapkan Turban, et al (2011) data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

2) Knowledge Discovery in Database (KDD)

Data mining merupakan salah satu bagian dari proses Knowledge Discovery in Databases (KDD). KDD merupakan proses mencari informasi yang lebih bernilai, lebih mudah dipahami dan baru dari penyimpanan data yang besar dan kompleks. Proses KDD menafsirkan hasil yang diperoleh dari sekumpulan data dengan menggabungkan dengan ilmu lainnya. Proses KDD dimulai dengan menetapkan tujuan dan diakhiri dengan evaluasi (Tomar & Agarwal, 2013). Tahapan dari KDD dapat dilihat pada Gambar 1 di bawah ini :



Gambar 1. Tahapan proses KDD

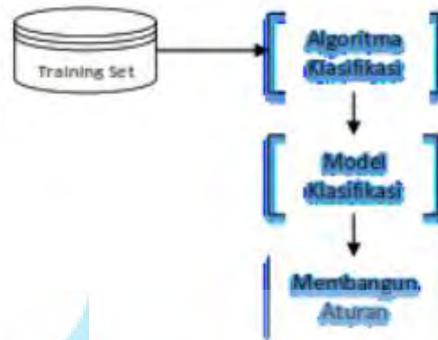
3) Klasifikasi

Algoritma data mining dapat dibagi menjadi tiga (Neelamegam & Ramaraj, 2013), yaitu supervised, unsupervised, dan semi-supervised. Dalam supervised learning, algoritma bekerja pada sekumpulan data yang telah diberi label atau telah diketahui kelasnya. Pada supervised learning, data belum diketahui label atau kelasnya, algoritma digunakan untuk mengelompokkan data berdasarkan kemiripannya. Sedangkan dalam semi supervised learning, sebagian kecil data telah memiliki label bersama dengan sejumlah data yang belum memiliki label. Klasifikasi termasuk ke dalam supervised learning.

Proses klasifikasi dibagi menjadi dua tahap (Annasaheb & Verma, 2016) yaitu :

1. Tahap membangun model

Pada langkah ini model klasifikasi dibangun berdasarkan data yang telah ditentukan kelasnya. Data sampel yang digunakan disebut sebagai data pelatihan atau data pembelajaran (training set). Proses ini disebut sebagai proses induksi yang ditunjukkan pada Gambar 2.



Gambar 2. Tahapan membangun model

2. Tahap menggunakan model klasifikasi

Pada tahap ini model diterapkan pada data yang belum diketahui kelasnya. Proses penerapan model klasifikasi untuk memprediksikan kelas label dari data dalam himpunan menggunakan data uji (testing set), proses ini disebut deduksi. Proses ini dapat dilihat pada Gambar. 3.



Gambar 3. Tahapan menggunakan model

4) Algoritma Klasifikasi

Dalam klasifikasi, terdapat variabel target yang bersifat kategoris yang dibagi menjadi kelas yang telah ditentukan, seperti kelas nasabah yang bermasalah atau tidak, hewan yang masuk ke dalam klasifikasi reptil, amfibi, mamalia, burung, atau ikan.

Model data mining memeriksa sejumlah besar data, setiap record berisi informasi tentang variabel target serta sekumpulan input atau prediktor variabel. Setiap algoritma klasifikasi yang digunakan akan menghasilkan model yang paling sesuai menghubungkan antara data input dan kelas klasifikasi yang telah diketahui sebelumnya.

Setiap algoritma bisa menghasilkan klasifikasi yang berbeda. Algoritma terbaik dapat dilihat dari data yang diklasifikasikan secara benar oleh model dengan data sebenarnya atau seberapa akurat model dapat memprediksi kelas klasifikasi.

Terdapat beberapa algoritma dan teknik yang digunakan pada data mining, diantara-Nya C4.5, Support Vector Machine, K-Nearest Neighbor, Naïve Bayes, Dan Artificial Neural Network (Nikam, 2015). Berikut ini penjelasan mengenai Naïve Bayes, dan K-Nearest Neighbor.

1. *Naïve Bayes*

Naïve Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema Bayes dengan asumsi independensi yang kuat. Dengan kata lain, dalam Naïve bayes menggunakan model fitur independen, maksud independen yang kuat pada fitur adalah bahwa data tidak berkaitan dengan data yang lain dalam kasus yang sama ataupun atribut yang lain. Persamaan dari teorema bayes adalah (Fadlan, Chairun., dkk. 2018):

$$P(X|Y) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Dimana :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)

$P(H)$: Probabilitas hipotesis H (prior Probabilitas)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

2. *K-Nearest Neighbors*

Algoritma ini merupakan pendekatan mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama. Disebut juga dengan pembelajar yang malas (lazy learners) karena hanya melihat kedekatan dengan tetangga (neighbor). Pada persamaan (4) merupakan salah satu rumus jarak yang digunakan dalam K Nearest Neighbor adalah jarak Euclidian (Gorunescu, 2011) :

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

Keterangan :

d_{ij} : jarak antara objek i dengan j

x_{ik} : nilai objek i pada variabel ke- k

x_{jk} : nilai objek j pada variabel ke- k

p : banyaknya variabel yang diamati

4) RapidMiner

Dalam pengolahan data mining umumnya digunakan software sebagai alat bantu. Beberapa software data mining diantaranya RapidMiner, weka, clementine, tanagra dan lain-lain. Menurut www.rapidminer.com, software rapidminer digunakan untuk merancang aliran secara visual untuk menganalisis data science dan machine learning di dalam tim mulai dari analis hingga pakar.

Rapidminer memiliki kemudahan dalam penggunaan, dapat mengumpulkan data dari semua sumber seperti basis data, cloud, dokumen, media sosial dan aplikasi bisnis. Selain itu dapat mengeksplorasi dan memvisualisasi data secara statistik. Tersedia beberapa model mesin pembelajaran dan model validasi.

II METODOLOGI PENELITIAN

II.I. Objek Penelitian

Tempat pelaksanaan penelitian ini di TK Santa Ursula Jakarta tepat di Jl. Pos. No.2, Pasar Baru, Sawah Besar, Jakarta Pusat.

1) Profil TK Santa Ursula Jakarta

TK Santa Ursula Jakarta yang berada di bawah naungan Yayasan Satya Bhakti kini terus berbenah dan mengembangkan diri. Berbagai pelatihan, baik bagi pendidik maupun peserta didik terus digalakkan. Berbagai model pembelajaran pun terus dipelajari dan dievaluasi. Sarana dan prasarana penunjang pembelajaran juga semakin ditingkatkan. Semua itu bertujuan untuk semakin meningkatkan mutu pelayanan pendidikan.

Pada tahun 2010 TK Santa Ursula menambah program layanan dengan menghadirkan Kelompok Bermain Santa Ursula Jakarta. Dengan demikian sejak tahun tersebut sekolah ini bernama TB – TK Santa Ursula Jakarta

2) Visi

Komunitas Pembelajar yang cerdas, beriman dan kasih.

3) Misi

- Menciptakan suasana yang kondusif bagi komunitas untuk

belajar terus-menerus.

- Mengembangkan semua potensi kecerdasan dengan memanfaatkan ilmu pengetahuan dan teknologi.
- Mengasah hati nurani sehingga anggota komunitas dapat hidup jujur, disiplin, dan bertanggung jawab.
- Mengembangkan religiusitas dan nilai-nilai kemanusiaan sehingga anggota komunitas dapat lebih menghayati imannya serta menghargai pluralitas dalam masyarakat.
- Menumbuhkembangkan sikap kasih dan peduli terhadap lingkungan dan sesama atas dasar kesetaraan gender dan semangat Serviam.
- Membekali dan mempersiapkan para siswa untuk melanjutkan pendidikan ke jenjang yang lebih tinggi.

II.II. Jenis Data

Jenis data yang digugkan dalam penelitian ini yaitu :

- 1) Data primer yaitu data yang diperoleh secara langsung dari objek penelitian yaitu : Laporan Perkembangan Peserta Didik TK Santa Ursula dengan jumlah murid TK B 60 siswa. Dari penelitian ini maka diperoleh data input pada gambar 4 yaitu sebagai berikut :

LAPORAN PERKEMBANGAN PESERTA DIDIK
TB - TK SANTA URSULA JAKARTA
Tahun Pelajaran 2019 - 2020


UNIVERSITAS
MERCU BUANA

Nama : [REDACTED]
Nomor Induk : [REDACTED]
Umur : 5 tahun
Kelas : [REDACTED]
Semester : 2

TB - TK SANTA URSULA
Jl. Pos No. 2 Jakarta Pusat 10010 Telp. 3846287
Website: www.tk.santaursulajakarta.kb.ji

Perkembangan Nilai Agama dan Moral

Capaian perkembangan [REDACTED] pada saat ini sesuai harapan, antara lain diuraikan dengan:

- Mendengarkan dengan sungguh-sungguh cerita kisah
- Menyebutkan contoh perbuatan menyayang ciptaan Tuhan
- Bertugas dalam badan dengan percaya diri dan sungguh-sungguh
- Menyebutkan pemimpin bangsa dan hari raya keagamaan
- Menyebutkan manfaat berbuat baik pada orang lain
- Menunjukkan pengalaman berbuat baik pada orang lain
- Melakukan 5S (Senyum, Sapa, Salam, Sapaan, Santun)
- Menyapa mulut apabila mengayuh, bertak, atau bertram
- Mengucapkan terima kasih kepada keluar kelas
- Diikuti di kursi dengan sikap yang baik
- Menggunakan pilihan kata yang baik dalam berbicara

Perkembangan Fisik - Motorik

Capaian perkembangan [REDACTED] saat ini sesuai harapan, diuraikan dengan:

- Menggerakkan setiap anggota tubuh
- Menyebutkan bagian anggota tubuh berdasarkan ciri/fungsinya
- Berjalan dengan satu kaki secara mandiri
- Merangkak dengan berbagai variasi
- Melampirkan benda pada sasaran dengan satu atau dua tangan
- Menghindarkan tempo sambil berjalan
- Memasukkan objek ke lubang kecil
- Melakukan gerakan koordinasi tangan dan kaki dengan cakatan
- Membuat tali kany dengan teknik memencik
- Melipat 5-6 lipatan membentuk sesuatu
- Memanca dengan manik-manik kecil
- Membuat tangkapan dan penangki dengan rapi
- Menggantung berbagai bahan sesuai pola
- Menjahit jalur dan silang dengan berbagai media
- Menunjukkan manfaat berolahraga
- Menyebutkan contoh olahraga yang sudah dilakukan

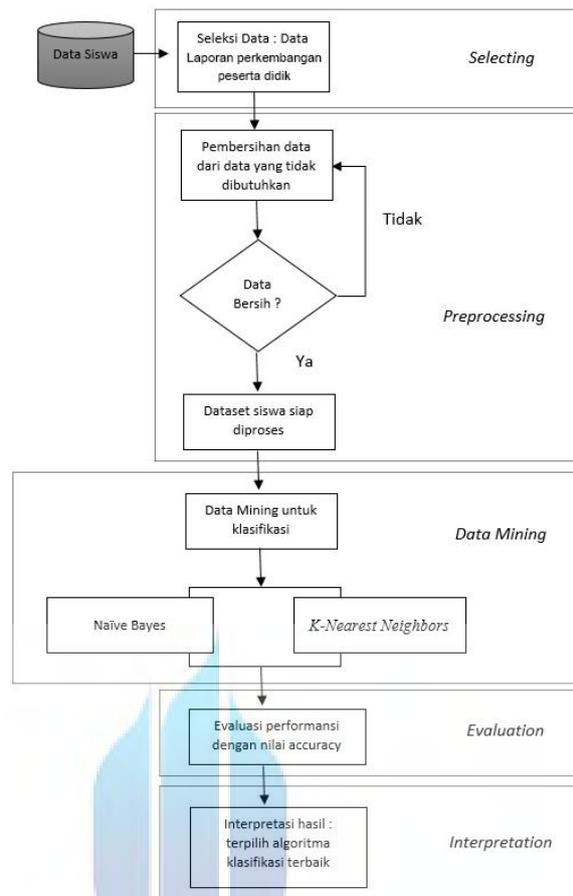
TB – TK SANTA URSULA		TB – TK SANTA URSULA	
Jl. Pos No. 2 Jakarta Pusat 10010 Telp. 3846287 Website : www.tb-tk.santasursulajakarta.sch.id		Jl. Pos No. 2 Jakarta Pusat 10010 Telp. 3846287 Website : www.tb-tk.santasursulajakarta.sch.id	
* * * * * • Menyebutkan manfaat telur yang cukup • Menceritakan pengalaman terlewat bangun • Menipiskan tempat makan dan area makannya • Melapas, memakai, dan melipat pakaian secara mandiri • Menata isi tas dengan rapi • BAK dan BAB sendiri dengan benar		* * * * * • Menunjukkan insiatif dalam kelompok • Menunjukkan kegigihan dalam mengerjakan tugas • Mau menerima kritik dan saran • Menceritakan peristiwa sesuai yang terjadi • Memperhatikan hak dan pendapatnya • Menunjukkan sikap penting menyuruh • Berani menelaah hak orang lain • Menjaga barang milik sendiri dan orang lain • Menggunakan air dengan hemat • Mampu memilih satu dari dua atau tiga pilihan • Memelihara lingkungan (tidak mengotori/mencuci) • Selalu berusaha tepat waktu • Mengurus diri sendiri tanpa bantuan • Membantu teman memecahkan masalah • Menasihati secara sederhana teman yang melakukan kesalahan	
Perkembangan Kognitif Capaian perkembangan ● hingga saat ini sesuai harapan, antara lain ditunjukkan dengan : • Membedakan konsep terapan, melayang, dan tenggelam melalui percobaan • Membedakan macam-macam suara • Mau bertanya bila menghadapi masalah • Membedakan berbagai macam bau melalui percobaan • Membilang dengan benda 1-20 • Memecahkan 5-7 kepingan gambar • Melakukan penjumlahan dan pengurangan 1-20 dengan benda • Membedakan konsep lebih banyak dan lebih sedikit dari dua kumpulan benda • Membedakan bentuk geometri berdasarkan cernya • Menceritakan tentang suatu informasi yang diperoleh dari buku • Mengelompokkan benda berdasarkan fungsinya • Membedakan konsep depan, belakang, dan tengah melalui kegiatan bermain • Membedakan konsep jauh dan dekat melalui kegiatan mengukur • Melakukan pengukuran dengan satuan langkah dan jagal • Menjelaskan konsep pertama sampai dengan ketujuh melalui kegiatan bermain • Menyebutkan alat transportasi darat, laut, dan udara • Membuat dan menggunakan telepon-telepon dari barang dan gelat plastik • Berkomunikasi melalui antaran • Menggunakan alat masak untuk membuat teh manis • Menggunakan serutan untuk merusut pensil		Perkembangan Bahasa Capaian perkembangan ● pada saat ini yang sesuai harapan, antara lain ditunjukkan dengan : • Membaca gambar dan buku cerita bergambar kemudian menceritakannya • Mengurutkan gambar menjadi gambar seri • Menjawab pertanyaan dengan tepat • Membuat gambar kemudian menceritakan maknanya • Membuat kalimat tanya dengan tepat • Membaca kata berdasarkan gambar atau lambang bilangan • Menyusun huruf dengan flash card menjadi kata sederhana • Menjelaskan ketertarikan terhadap buku/media cetak • Menuliskan nama sendiri dengan tepat • Menulis kata sederhana • Menulis nama bilangan 1-10 • Menulis kalimat pendek yang terdiri atas dua kata sederhana Hal yang cukup berkembang dan perlu ditingkatkan terus adalah : • Berbicara dan berbicara lancar lebih bernilai lagi	
Perkembangan Sosial-Emosional Capaian perkembangan ● pada saat ini sesuai harapan, antara lain ditunjukkan dengan : • Mengikuti aturan permainan/kegiatan		Perkembangan Seni Capaian perkembangan ● pada saat ini yang sangat baik, antara lain ditunjukkan dengan : • Menggambar orang dengan langkah dan proporsional • Membuat gambar dengan teknik kalase • Menggambar bebas dari bentuk dasar lingkaran/persegi Capaian perkembangan ● pada saat ini yang sesuai harapan, antara lain ditunjukkan dengan : • Menyebutkan contoh lingkungan sosial yang pernah dijumpai • Menyebutkan contoh lagu daerah • Menggambar orang dengan langkah dan proporsional • Membuat gambar dengan teknik kalase • Mengucapkan sajak tentang lingkungan alam • Menyanyi lagu anak • Menjelaskan ciri-ciri cuaca tertentu • Menari sesuai irama musik • Mencocok sesuai pola • Melakukan senam irama dengan sungguh-sungguh • Menceritakan gerakan pantomim ke dalam bahasa lisan • Menganyam dengan berbagai media • Membuat karya dari barang bekas • Membuat alat musik perkusi sederhana dan memainkannya secara berirama • Menciptakan bentuk yang lebih kompleks dari balok	

Gambar 4 Sampel Data Primer

2) Data sekunder yaitu data yang diperoleh dari jurnal, literature, buku referensi.

II.III. Metode Penelitian

Model yang digunakan dalam penelitian ini mengikuti tahapan yang ada pada proses *Knowledge discovery in database* (KDD). Pada Gambar 5 adalah penjelasan untuk masing-masing tahap yang akan dilakukan.



Gambar 5. Tahapan penelitian berdasarkan KDD
Berikut ini penjelasan tiap tahap penelitian pada gambar 5 :

- 1) Selection
Pada tahap ini dilakukan seleksi data laporan perkembangan peserta didik yang terdiri dari variabel-variabel predictor dan satu target variabel. Variabel target yaitu klasifikasi kesiapan anak menuju ke jenjang sekolah dasar. Sedangkan variabel-variabel predictor yaitu, jenis kelamin, usia, indikator perkembangan.
- 2) Preprocessing
Data yang diambil sesuai dengan laporan perkembangan peserta didik di TK Santa Ursula. Dari data yang diambil dilakukan pembersihan data yang tidak diperlukan, data yang hilang atau data ganda.
- 3) Data Mining
Pada tahap ini dilakukan pemilihan teknik data mining yang sesuai. Untuk fungsi klasifikasi digunakan algoritma *K-Nearest Neighbors* dan *Naïve Bayes*. Karena klasifikasi merupakan *supervised learning* maka berikut ini adalah tahapan dalam model *supervised learning* (Larose, 2005)
- 4) Evaluation
Tahap ini digunakan untuk mengevaluasi hasil-hasil prediksi yang dihasilkan oleh kedua algoritma dan dipilih metode

algoritma yang menghasilkan nilai mendekati klasifikasi data sebenarnya. Evaluasi dilakukan dengan menggunakan metode *Confusion Matrix*. Nilai performansi yang digunakan yaitu accuracy dan error.

III HASIL DAN PEMBAHASAN

III.I Hasil Pengujian

III.I.I Data Selection

Dalam menganalisa kesiapan anak untuk menyesuaikan diri dijenjang sekolah dasar ada beberapa atribut dataset yang digunakan, dataset yang digunakan adalah sebagai berikut :

Gender	1) Pria : 45 Orang 2) Perempuan : 63 Orang	IFM20	Sesuai Harapan : 108
Usia	1) 5 Tahun : 16 Orang 2) 6 Tahun : 92 Orang	IFM21	Sesuai Harapan : 108
IA1	Sesuai Harapan : 108	IFM22	Sesuai Harapan : 108
IA2	Sesuai Harapan : 108	IK1	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA3	Sesuai Harapan : 108	IK2	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA4	Sesuai Harapan : 108	IK3	1) Sesuai Harapan : 104 2) Perlu Ditingkatkan : 3 3) Sangat Baik : 1
IA5	Sesuai Harapan : 108	IK4	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA6	Sesuai Harapan : 108	IK5	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA7	1) Sesuai Harapan : 102 2) Perlu Ditingkatkan : 6	IK6	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA8	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1	IK7	1) Sesuai Harapan : 103 2) Perlu Ditingkatkan : 4 3) Sangat Baik : 1
IA9	Sesuai Harapan : 108	IK8	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 1 3) Sangat Baik : 1
IA10	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2	IK9	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IA11	Sesuai Harapan : 108	IK10	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM1	Sesuai Harapan : 108	IK11	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM2	Sesuai Harapan : 108	IK12	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM3	Sesuai Harapan : 108	IK13	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM4	Sesuai Harapan : 108	IK14	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM5	Sesuai Harapan : 108	IK15	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM6	Sesuai Harapan : 108	IK16	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM7	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2	IK17	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM8	Sesuai Harapan : 108	IK18	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IFM9	Sesuai Harapan : 108	IK19	1) Sesuai Harapan : 107
IFM10	1) Sesuai Harapan : 104 2) Perlu Ditingkatkan : 4		
IFM11	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1		
IFM12	Sesuai Harapan : 108		
IFM13	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1		
IFM14	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2		
IFM15	Sesuai Harapan : 108		
IFM16	Sesuai Harapan : 108		
IFM17	Sesuai Harapan : 108		
IFM18	Sesuai Harapan : 108		
IFM19	Sesuai Harapan : 108		

	2) Sangat Baik : 1
IK20	1) Sesuai Harapan : 107 2) Sangat Baik : 1
ISE1	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2
ISE2	1) Sesuai Harapan : 101 2) Perlu Ditingkatkan : 7
ISE3	1) Sesuai Harapan : 103 2) Perlu Ditingkatkan : 5
ISE4	Sesuai Harapan : 108
ISE5	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1
ISE6	Sesuai Harapan : 108
ISE7	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2
ISE8	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1
ISE9	1) Sesuai Harapan : 105 2) Perlu Ditingkatkan : 3
ISE10	Sesuai Harapan : 108
ISE11	Sesuai Harapan : 108
ISE12	1) Sesuai Harapan : 107 2) Perlu Ditingkatkan : 1
ISE13	1) Sesuai Harapan : 100 2) Perlu Ditingkatkan : 8
ISE14	Sesuai Harapan : 108
ISE15	1) Sesuai Harapan : 106 2) Perlu Ditingkatkan : 2
ISE16	1) Sesuai Harapan : 105 2) Perlu Ditingkatkan : 3
IB1	1) Perlu Ditingkatkan: 3 2) Sesuai Harapan : 101 3) Sangat Baik : 4
IB2	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IB3	1) Perlu Ditingkatkan: 1 2) Sesuai Harapan : 103 3) Sangat Baik : 4
IB4	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IB5	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IB6	1) Perlu Ditingkatkan: 3 2) Sesuai Harapan : 99 3) Sangat Baik : 6
IB7	1) Perlu Ditingkatkan: 4 2) Sesuai Harapan : 102 3) Sangat Baik : 2

IB8	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IB9	1) Sesuai Harapan : 107 2) Sangat Baik : 1
IB10	1) Sesuai Harapan : 104 2) Sangat Baik : 4
IB11	1) Perlu Ditingkatkan: 1 2) Sesuai Harapan : 105 3) Sangat Baik : 2
IB12	1) Perlu Ditingkatkan: 5 2) Sesuai Harapan : 101 3) Sangat Baik : 2
IB13	1) Sesuai Harapan : 106 2) Sangat Baik : 2
IB14	1) Perlu Ditingkatkan: 6 2) Sesuai Harapan : 100 3) Sangat Baik : 2
IS1	Sesuai Harapan : 108
IS2	Sesuai Harapan : 108
IS3	1) Perlu Ditingkatkan: 1 2) Sesuai Harapan : 105 3) Sangat Baik : 2
IS4	1) Perlu Ditingkatkan: 4 2) Sesuai Harapan : 102 3) Sangat Baik : 2
IS5	1) Perlu Ditingkatkan: 3 2) Sesuai Harapan : 101 3) Sangat Baik : 4
IS6	1) Sesuai Harapan : 96 2) Perlu Ditingkatkan : 12
IS7	1) Perlu Ditingkatkan: 5 2) Sesuai Harapan : 101 3) Sangat Baik : 2
IS8	Sesuai Harapan : 108
IS9	1) Perlu Ditingkatkan: 2 2) Sesuai Harapan : 104 3) Sangat Baik : 2
IS10	1) Perlu Ditingkatkan: 1 2) Sesuai Harapan : 107
IS11	Sesuai Harapan : 108
IS12	Sesuai Harapan : 108
IS13	1) Perlu Ditingkatkan: 2 2) Sesuai Harapan : 106
IS14	1) Sesuai Harapan : 106 2) Sangat Baik: 2
IS15	Sesuai Harapan : 108
IS16	1) Sesuai Harapan : 106 2) Sangat Baik: 2

Keterangan :

Keterangan	Kode
Mendengarkan dengan sungguh-sungguh keagamaan dengan cerita	IA1
Menyebutkan pemimpin ibadah dan hari raya keagamaan	IA2
Menyebutkan perbuatan menyayangi ciptaan Tuhan contoh	IA3

Bertugas dalam ibadah dengan percaya diri dan sungguh-sungguh	IA4
Menyebutkan manfaat berbuat baik pada orang lain	IA5
Menceritakan pengalaman berbuat baik pada orang lain	IA6
Melakukan 5S (Senyum, Sapa, Salam, Sopan, Santun)	IA7
Menutup mulut apabila menguap, batuk, atau bersin	IA8

Mengucapkan permisi jika hendak keluar kelas	IA9
Duduk di kursi dengan sikap yang baik	IA10
Menggunakan pilihan kata yang baik dalam berbicara	IA11
Menggerakkan setiap anggota tubuh	IFM1
Mengidentifikasi anggota tubuh berdasarkan ciri/fungsinya	IFM2
Berdiri dengan satu kaki secara seimbang	IFM3
Merangkak dengan berbagai variasi	IFM4
Melemparkan benda pada sasaran dengan satu atau dua tangan	IFM5
Mengelindungi sampai sambil berjalan	IFM6
Memasukkan objek ke lubang kecil	IFM7
Melakukan gerakan koordinasi tangan dan kaki dengan cekatan	IFM8
Membuat hasil karya dengan teknik memercik	IFM9
Melipat 5-6 lipatan membentuk sesuatu	IFM10
Meronce dengan manik-manik kecil	IFM11
Membuat lingkaran dan persegi dengan rapi	IFM12
Menggunting berbagai bahan sesuai pola	IFM13
Menjahit jelujur dan silang dengan berbagai media	IFM14
Menjelaskan manfaat berolahraga	IFM15
Menyebutkan contoh olahraga yang mudah dilakukan	IFM16
Menyebutkan manfaat tidur yang cukup	IFM17
Menceritakan pengalaman terlambat bangun	IFM18
Merapikan tempat makan dan area makannya	IFM19
Melepas, memakai, dan melipat pakaian secara mandiri	IFM20
Menata isi tas dengan rapi	IFM21
BAK dan BAB sendiri dengan benar	IFM22
Membedakan konsep terapung, melayang, dan tenggelam melalui percobaan	IK1
Membedakan macam-macam suara	IK2
Mau bertanya bila menghadapi masalah	IK3
Membedakan berbagai macam bau melalui percobaan	IK4

Membilang dengan benda 1-20	IK5
Mencari 5-7 kejanggalan gambar	IK6
Melakukan penjumlahan dan pengurangan 1-20 dengan benda	IK7
Membedakan konsep lebih banyak dan lebih sedikit dari dua kumpulan benda	IK8
Membedakan bentuk geometri berdasarkan cirinya	IK9
Menceritakan tentang suatu informasi yang diperoleh dari buku	IK10
Mengelompokkan benda berdasarkan fungsinya	IK11
Membedakan konsep depan, belakang, dan tengah melalui kegiatan bermain	IK12
Membedakan konsep jauh dan dekat melalui kegiatan mengukur	IK13
Melakukan pengukuran dengan satuan langkah dan jengkal	IK14
Menjelaskan konsep pertama sampai dengan ketujuh melalui kegiatan bermain	IK15
Menyebutkan alat transportasi darat, laut, dan udara	IK16
Membuat dan menggunakan telepon-teleponan dari benang dan gelas plastik	IK17
Berkomunikasi melalui interkom	IK18
Menggunakan alat memasak untuk membuat teh manis	IK19
Menggunakan serutan untuk meraut pensil	IK20
Mengikuti aturan permainan/kegiatan	ISE1
Menunjukkan inisiatif dalam kelompok	ISE2
Menunjukkan kegigihan dalam mengerjakan tugas	ISE3
Mau menerima kritik dan saran	ISE4
Menceritakan peristiwa sesuai yang terjadi	ISE5
Mempertahankan hak dan pendapatnya	ISE6
Menunjukkan sikap pantang menyerah	ISE7
Berani membela hak orang lain	ISE8
Menjaga barang milik sendiri dan orang lain	ISE9
Menggunakan air dengan hemat	ISE10
Mampu memilih satu dari dua atau tiga pilihan	ISE11

Memelihara lingkungan (tidak mengotori/merusak)	ISE12
Selalu berusaha tepat waktu	ISE13
Mengurus diri sendiri tanpa bantuan	ISE14
Membantu teman memecahkan masalah	ISE15
Menasihati secara sederhana teman yang melakukan kesalahan	ISE16
Membaca gambar dan buku cerita bergambar kemudian menceritakannya	IB1
Mengurutkan gambar menjadi gambar seri	IB2
Menjawab pertanyaan dengan tepat	IB3
Membuat gambar kemudian menceritakan maksudnya	IB4
Membuat kalimat tanya dengan tepat	IB5
Berbicara lancar dengan kalimat yang lebih kompleks (5-6 kata)	IB6
Bercerita menggunakan kata ganti orang	IB7
Membaca kata berdasarkan gambar atau lambang bilangan	IB8
Menyusun huruf dengan flash card menjadi kata sederhana	IB9
Menunjukkan ketertarikan terhadap buku/media cetak	IB10
Menuliskan nama sendiri dengan tepat	IB11
Menulis kata sederhana	IB12
Menulis nama bilangan 1-10	IB13
Menulis kalimat pendek yang terdiri atas dua kata sederhana	IB14
Menyebutkan contoh lingkungan sosial yang pernah dijumpai	IS1
Menyebutkan contoh lagu daerah	IS2
Menggambar bebas dari bentuk dasar lingkaran/persegi	IS3
Menggambar orang dengan lengkap dan proporsional	IS4
Membuat gambar dengan teknik kolase	IS5
Mengucapkan sajak tentang lingkungan alam	IS6
Menyanyi lagu anak	IS7
Menjelaskan ciri-ciri cuaca tertentu	IS8
Menari sesuai irama musik	IS9
Mencocok sesuai pola	IS10

Melakukan senam irama dengan sungguh-sungguh	IS11
Menceritakan gerakan pantomim ke dalam bahasa lisan	IS12
Menganyam dengan berbagai media	IS13
Membuat karya dari barang bekas	IS14
Membuat alat musik perkusi sederhana dan memainkannya secara berirama	IS15
Menciptakan bentuk yang lebih kompleks dari balok	IS16

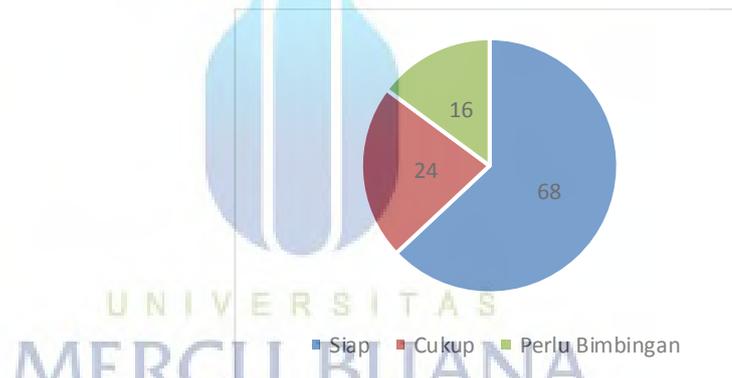
III.I.II Preprocessing Data

Untuk memperoleh dataset untuk prediksi kesiapan anak menuju ke jenjang sekolah dasar diperlukan pembersihan data. Contoh pengumpulan data disusun pada gambar 6 dibawah ini :

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	NAMA	ORANG tua	JKL	IAI	IEE	IGM	IKS	IKL	IKP	IKM	IKN	IKO	IKP	IKQ	IKR
2	1	FRAN	♀	TANJUNGPURA	SEKOLAH										
3	2	FRAN	♀	TANJUNGPURA	SEKOLAH										
4	3	FRAN	♀	TANJUNGPURA	SEKOLAH										
5	4	FRAN	♀	TANJUNGPURA	SEKOLAH										
6	5	FRAN	♀	TANJUNGPURA	SEKOLAH										
7	6	FRAN	♀	TANJUNGPURA	SEKOLAH										
8	7	FRAN	♀	TANJUNGPURA	SEKOLAH										
9	8	FRAN	♀	TANJUNGPURA	SEKOLAH										
10	9	FRAN	♀	TANJUNGPURA	SEKOLAH										
11	10	FRAN	♀	TANJUNGPURA	SEKOLAH										
12	11	FRAN	♀	TANJUNGPURA	SEKOLAH										
13	12	FRAN	♀	TANJUNGPURA	SEKOLAH										
14	13	FRAN	♀	TANJUNGPURA	SEKOLAH										
15	14	FRAN	♀	TANJUNGPURA	SEKOLAH										
16	15	FRAN	♀	TANJUNGPURA	SEKOLAH										
17	16	FRAN	♀	TANJUNGPURA	SEKOLAH										
18	17	FRAN	♀	TANJUNGPURA	SEKOLAH										
19	18	FRAN	♀	TANJUNGPURA	SEKOLAH										
20	19	FRAN	♀	TANJUNGPURA	SEKOLAH										
21	20	FRAN	♀	TANJUNGPURA	SEKOLAH										
22	21	FRAN	♀	TANJUNGPURA	SEKOLAH										
23	22	FRAN	♀	TANJUNGPURA	SEKOLAH										
24	23	FRAN	♀	TANJUNGPURA	SEKOLAH										
25	24	FRAN	♀	TANJUNGPURA	SEKOLAH										
26	25	FRAN	♀	TANJUNGPURA	SEKOLAH										
27	26	FRAN	♀	TANJUNGPURA	SEKOLAH										
28	27	FRAN	♀	TANJUNGPURA	SEKOLAH										
29	28	FRAN	♀	TANJUNGPURA	SEKOLAH										
30	29	FRAN	♀	TANJUNGPURA	SEKOLAH										

Gambar 6. Data yang telah dilakukan pembersihan.

Berdasarkan data Laporan Perkembangan Peserta didik TK Santa Ursula yang Siap menuju ke jenjang sekolah dasar sebanyak 68 Dan dapat dilihat pada gambar 7 berikut ini :

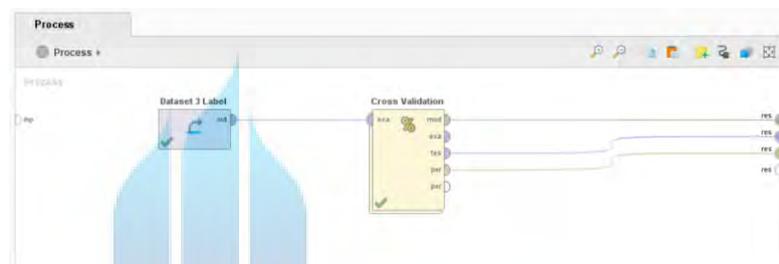


Gambar 7. Perbandingan Klasifikasi Kesiapan “Perlu Bimbingan” , “Cukup” , “Siap”.

III.I.III Data Mining

Pada pengolahan data, dilakukan tahap pemodelan untuk proses klasifikasi yaitu menerapkan algoritma *Naive Bayes* dan *K-Nearest Neighbor* Pengolahan data menggunakan perangkat lunak RapidMiner Version 9.8.001. Jumlah data yang akan diolah sebanyak 108 data. Proses pengambilan data pada RapidMiner untuk algoritma *Naive Bayes* dan *K-Nearest Neighbor* langsung pada format excel dari data yang telah dipreprocessing seperti gambar 8 dibawah. Selanjutnya dilakukan *Select Attributes* untuk dilakukan pembersihan data yang tidak digunakan / berpengaruh seperti nama. Selanjutnya dilakukan *Cross Validation* untuk data yang telah diambil. Teknik validasi yang digunakan pada proses klasifikasi adalah *K-Fold Cross Validation*

K-Fold Cross Validation adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model klasifikasi dimana data dipisahkan menjadi dua bagian yaitu data proses latih (training) dan data uji, *K-Fold Cross Validation* digunakan karena dapat mengurangi waktu komputasi dengan tetap menjaga keakuratan estimasi. Nilai k diambil 6 Fold sehingga dari 108 data akan menjadi subset data dengan ukuran sama yaitu sekitar 18 data. Dari masing-masing 6 subset tersebut, 90 data menjadi data latih dan 18 data menjadi data uji untuk algoritma *K-Nearest Neighbor* sedangkan *Naïve Bayes* Nilai k diambil 3 Fold sehingga dari 108 data akan menjadi subset data dengan ukuran sama yaitu sekitar 36 data. Dari masing-masing 3 subset tersebut, 72 data menjadi data latih dan 36 data menjadi data uji.

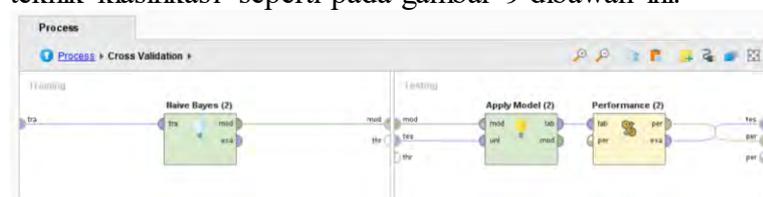


Gambar 8. Proses pengambilan data untuk algoritma *Naïve Bayes* dan *K-Nearest Neighbor*

Pada proses validasi dilakukan penerapan algoritma-algoritma data mining sesuai dengan yang digunakan pada penelitian ini yaitu algoritma *Naïve Bayes* dan *K-Nearest Neighbor*

III.I.IV Hasil Perhitungan Klasifikasi dengan RapidMiner

Pada data training diterapkan algoritma *Naïve Bayes* untuk teknik klasifikasi seperti pada gambar 9 dibawah ini:



Gambar 9. Model Klasifikasi dengan algoritma *Naïve Bayes*

Output hasil performasi dari algoritma ini berupa pengklasifikasi kesiapan yang termasuk ke dalam klasifikasi kesiapan “Perlu Bimbingan” , “Cukup” dan “Siap”. Yang diterapkan pada data testing.

Jumlah data yang diprediksi dengan benar oleh algoritma *Naïve Bayes* ditunjukkan dalam Tabel 1. *Confusion Matrix* berikut :

Tabel 1. Confusion Matrix Algoritma Naïve Bayes

	True Perlu Bimbingan	True Cukup	True Siap	Class Precision
Pred. Perlu Bimbingan	2	3	5	20.00%
Pred. Cukup	4	5	11	25.00%
Pred. Siap	10	16	52	66.67%
Class Recall	12.50%	20.83%	76.47%	

Keterangan table 5 adalah :

- Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi PERLU BIMBINGAN adalah 2.
- Jumlah data sebenarnya yang CUKUP dan diprediksi CUKUP adalah 5.
- Jumlah data sebenarnya yang SIAP dan diprediksi SIAP adalah 52.
- Jumlah data sebenarnya yang CUKUP dan diprediksi PERLU BIMBINGAN adalah 3.
- Jumlah data sebenarnya yang SIAP dan diprediksi PERLU BIMBINGAN adalah 5.
- Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi CUKUP adalah 4.
- Jumlah data sebenarnya yang SIAP dan diprediksi CUKUP adalah 11.
- Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi SIAP adalah 10.
- Jumlah data sebenarnya yang CUKUP dan diprediksi SIAP adalah 16.

Akurasi dari model yaitu :

$$\text{Akurasi} = \frac{52+5+2}{2+5+52+3+5+4+11+10+16} = 54.63\%$$

$$\text{Error} = \frac{3+5+4+11+10+16}{2+5+52+3+5+4+11+10+16} = 45.37\%$$

Pada data Training diterapkan algoritma *K-Nearest Neighbor* untuk teknik klasifikasi seperti pada Gambar 10 dibawah ini. Nilai k yang diambil sebanyak 10 karena memberikan hasil performansi yang lebih baik dibandingkan dengan k = 1, 2, 3, 4, 5, 6, 7, 8, dan 9.



Gambar 10. Model Klasifikasi dengan algoritma *K-Nearest Neighbor*

Jumlah data yang diprediksi dengan benar oleh algoritma *K-Nearest Neighbor* ditunjukkan dalam *Confusion Matrix* seperti yang ditunjukkan pada Tabel 2. berikut ini :

Tabel 2. Confusion Matrix Algoritma *K-Nearest Neighbor*

	True Perlu Bimbingan	True Cukup	True Siap	Class Precision
Pred. Perlu Bimbingan	0	0	0	0.00%
Pred. Cukup	0	0	0	0.00%
Pred. Siap	16	24	68	62.96%
Class Recall	0.00%	0.00%	100.00%	

Keterangan table 5 adalah :

- a. Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi PERLU BIMBINGAN adalah 0.
- b. Jumlah data sebenarnya yang CUKUP dan diprediksi CUKUP adalah 0.
- c. Jumlah data sebenarnya yang SIAP dan diprediksi SIAP adalah 68.
- d. Jumlah data sebenarnya yang CUKUP dan diprediksi PERLU BIMBINGAN adalah 0.
- e. Jumlah data sebenarnya yang SIAP dan diprediksi PERLU BIMBINGAN adalah 0.
- f. Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi CUKUP adalah 0.
- g. Jumlah data sebenarnya yang SIAP dan diprediksi CUKUP adalah 0.
- h. Jumlah data sebenarnya yang PERLU BIMBINGAN dan diprediksi SIAP adalah 16.
- i. Jumlah data sebenarnya yang CUKUP dan diprediksi SIAP adalah 24.

Akurasi dari model yaitu :

$$\text{Akurasi} = \frac{0+0+68}{0+0+68+0+0+0+0+16+24} = 62.96\%$$

$$\text{Error} = \frac{0+0+0+0+16+24}{0+0+68+0+0+0+0+16+24} = 38.04\%$$

IV PENUTUP

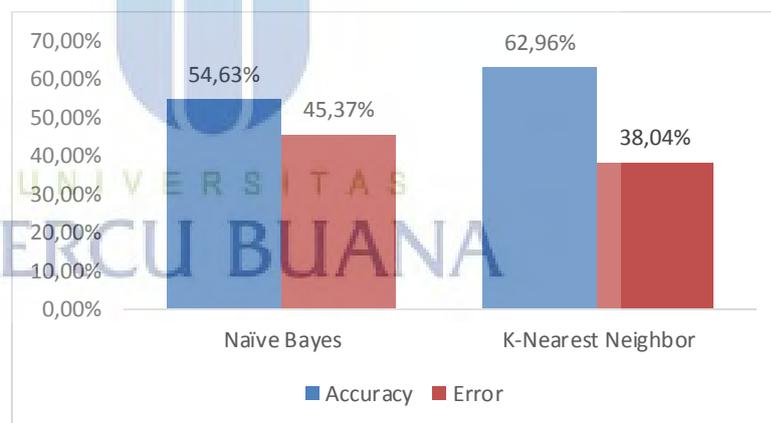
1. Kesimpulan

Hasil performansi pada setiap model yaitu accuracy dan error selanjutnya akan dibandingkan untuk mengetahui algoritma mana yang lebih baik dalam memprediksi kesiapan anak menuju ke jenjang sekolah dasar. Tabel 3. Berikut merupakan table perbandingan performansi *Naïve Bayes* dan *K-Nearest Neighbor*

Tabel 3. Perbandingan Nilai Performansi *Naïve Bayes* dan *K-Nearest Neighbor*

	Accuracy	Error
<i>Naïve Bayes</i>	54.63%	45.37%
<i>K-Nearest Neighbor</i>	62.96%	38.04%

Dari hasil perbandingan terlihat bahwa algoritma *K-Nearest Neighbor* memiliki nilai yang paling baik untuk kategori performansi dibandingkan dengan *Naïve Bayes*. Untuk Nilai Accuracy nilai terbesar adalah yang terbaik, sedangkan untuk error adalah nilai yang tekecil. Grafik perbandingan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* adalah sebagai berikut :



Dari hasil penelitian dapat disimpulkan bahwa :

1. Dengan teknik data mining dapat diperoleh informasi-informasi yang terdapat di dalam database peserta didik seperti pola kesiapan.
2. Dari algoritma yang diuji dapat digunakan untuk memprediksi kesiapan anak ke jenjang sekolah dasar, dilihat dari nilai accuracy algoritma *Naïve Bayes* dan *K-Nearest Neighbor* terdapat dalam kategori “Perlu Bimbingan” , “Cukup” dan “Siap”
3. Dari hasil evaluasi diperoleh hasil bahwa *K-Nearest Neighbor* yang paling baik untuk memprediksi kesiapan

anak menuju ke jenjang sekolah dasar karena memiliki nilai akurasi tertinggi dan error terkecil dibandingkan algoritma *Naïve Bayes*

2. Saran

Berikut ini adalah beberapa saran yang dapat dilakukan untuk penelitian selanjutnya :

- 1) Untuk meningkatkan nilai akurasi dari algoritma dapat ditambahkan data Laporan Perkembangan Peserta didik yang diperkirakan mempengaruhi tingkat kesiapan anak menuju kejenjang sekolah dasar
- 2) Dapat dilakukan pengujian algoritma lain selain algoritma yang digunakan dalam penelitian ini

V DAFTAR PUSTAKA

- Intan, Indo. 2019. "Sistem Prediksi Pre Test Penentuan Kesiapan Memasuki Pendidikan Dasar Pre Test Prediction System for Preparing Readiness for Basic Education." 119–24.
- Putri, Riyan Eko, Suparti, and Rita Rahmawati. 2014. "Perbandingan Metode Klasifikasi Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Data Status Kerja Di Kabupaten Demak Tahun 2012." *Jurnal Gaussian* 3(4):831–38.
- Rahman, Maulana Aditya, Nurul Hidayat, and Ahmad Afif Supianto. 2018. "Komparasi Metode Data Mining K-Nearest Neighbor Dengan Naïve Bayes Untuk Klasifikasi Kualitas Air Bersih (Studi Kasus PDAM Tirta Kencana Kabupaten Jombang)." *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (JPTIIK)* 2(12):925–28.
- Widaningsih, Sri. 2019. "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm." *Jurnal Tekno Insentif* 13(1):16–25.

KERTAS KERJA

Ringkasan

Hasil dari pengujian pengaruh nilai K-Fold menghasilkan akurasi naïve bayes optimum senilai 54.63% +/- 4.24% (micro average : 54.63%) pada nilai K-Fold 3, sedangkan pengujian nilai akurasi K-Nearest Neighbors optimum senilai 62.96% +/- 2.87% (micro average : 62.96%) dengan nilai K=10 pada nilai K-Fold 6. Dari hasil tersebut, disimpulkan bahwa metode K-Nearest Neighbors dapat digunakan dengan baik untuk memprediksi kesiapan anak masuk sekolah dasar.

Rasio perbandingan jumlah data *training* dan data *testing* juga memiliki pengaruh terhadap akurasi algoritme semakin besar selisih persentase atau rasio antara data *training* dan data *testing* maka semakin tinggi pula akurasi yang didapatkan. Hal tersebut dikarenakan jumlah data *training* yang semakin lebih banyak daripada jumlah data *testing* yang semakin lebih sedikit maka model classifier yang dibangun berdasarkan faktar dari data *training* akan lebih baik dan lebih lengkap untuk melakukan prediksi terhadap klasifikasi data baru atau data *testing*. Sehingga akurasi yang didapatkan jauh lebih baik juga

Hasil prediksi ini akan dibuktikan dengan hasil perkembangan di sekolah dasar Santa Ursula, apakah hasil prediksi berbeda atau sama dengan hasil perkembangan di sekolah dasar Santa Ursula.

UNIVERSITAS
MERCU BUANA