

# Metode *Naive Bayes Classifier* – *Smoothing* pada Sensor *Smartphone* untuk Klasifikasi Aktivitas Pengendara

## (*Naive Bayes Classifier* – *Smoothing Method on Smartphone Sensors for Driver Activity Classification*)

Haniah Mahmudah<sup>1</sup>, Okkie Puspitorini<sup>1</sup>, Nur Adi Siswandari<sup>1</sup>, Ari Wijayanti<sup>1</sup>, Eliya Alfatekha<sup>1</sup>

**Abstract**—Most causes of death are traffic accidents. This paper aims to obtain parameters of identification of driving activities that can be developed to detect accidents in the next studies. Data is collected by sensors on smartphones, using accelerometer and gyroscope sensors. The proposed method uses *Naive Bayes Classifiers* (NBC) algorithm to determine driving activity, by dividing dataset into training and testing data using *k-fold* parameters. NBC can work using less training data, by calculating the probability value of each class from means and variance of each feature to classify classes efficiently. The results show that the accuracy of the classification is higher if a smoothing process is carried out, using single exponential smoothing method, before the classification process of the NBC algorithm is done. The testing using 8 *k-fold* CV without smoothing process, using smoothing alpha ( $\alpha$ ) = 0.1, and using  $\alpha$  = 0.9 obtain the accuracy of 98.43%, 99.27%, and 98.43%, respectively. It can be concluded that the NBC method combined with smoothing method using  $\alpha$  = 0.1 produces greater accuracy.

**Intisari**—Sebagian besar penyebab kematian diakibatkan oleh kecelakaan lalu lintas. Makalah ini bertujuan mendapatkan parameter identifikasi aktivitas berkendara yang dapat dikembangkan untuk mendeteksi kecelakaan pada penelitian selanjutnya. Data dikumpulkan dengan memanfaatkan sensor pada *smartphone*, menggunakan sensor akselerometer dan giroskop. Metode statistik yang diusulkan menggunakan algoritme *Naive Bayes Classifiers* (NBC) untuk menentukan aktivitas berkendara, dengan membagi *dataset* ke dalam data latih dan data uji menggunakan beberapa parameter *k-fold*. Algoritme NBC dapat bekerja dengan sedikit data latih, dengan cara menghitung nilai probabilitas setiap kelas dari perhitungan jumlah rata-rata dan varians setiap fitur untuk mengklasifikasikan kelas secara linier dengan sangat efisien. Hasil menunjukkan bahwa akurasi klasifikasi aktivitas berkendara lebih tinggi apabila terdapat proses *smoothing* menggunakan metode *single exponential smoothing* sebelum proses klasifikasi algoritme NBC. Hasil pengujian menggunakan 8 *k-fold* CV tanpa proses *smoothing*, dengan *smoothing alpha* ( $\alpha$ ) = 0,1, dan dengan  $\alpha$  = 0,9 mendapatkan akurasi, berturut-turut, sebesar 98,43%, 99,27%, dan 98,43%. Hasil komparasi metode klasifikasi aktivitas berkendara pada makalah ini membuktikan bahwa metode NBC yang dikombinasikan dengan metode *smoothing*  $\alpha$  = 0,1 menghasilkan akurasi yang lebih besar.

**Kata Kunci**—*Naive Bayes*, *Smoothing*, Giroskop, Akselerometer, *Smartphone*.

### I. PENDAHULUAN

Kecelakaan lalu lintas dapat terjadi di mana saja, termasuk di lokasi dengan intensitas kendaraan yang rendah atau jalur yang jarang dilewati kendaraan, seperti jalanan pegunungan. Menurut data laporan kecelakaan lalu lintas secara global tahun 2016 oleh *World Health Organization* (WHO), kecelakaan lalu lintas termasuk dalam sepuluh besar penyebab kematian di seluruh dunia, menewaskan lebih dari 1,35 juta orang per tahun [1]. Berdasarkan data Korlantas Polri, pada tahun 2017 hingga 2018, total kasus kecelakaan di Indonesia mencapai sekitar 25.000 kasus, dengan 5.000 kasus mengakibatkan kematian [2].

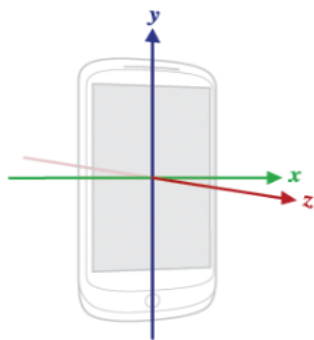
Seiring perkembangan teknologi yang pesat, beberapa penelitian telah dilakukan untuk mengatasi masalah kecelakaan, salah satunya menggunakan *smartphone* berbasis Android dengan memanfaatkan sensor giroskop dan akselerometer. Sensor tersebut berfungsi mengumpulkan data yang mendeteksi aktivitas berkendara seperti mesin menyala, berjalan lurus, belok kiri, dan belok kanan. Selain itu, sensor akselerometer dan giroskop digunakan untuk mendapatkan nilai percepatan terhadap tiga koordinat sumbu x, y, dan z, yang merepresentasikan kondisi kemiringan kendaraan sebagai parameter deteksi kecelakaan.

Menurut hasil penelitian mengenai perbandingan beberapa kinerja algoritme, disimpulkan bahwa prediksi mode kendaraan menggunakan *Naive Bayes Classifiers* (NBC) menghasilkan akurasi yang lebih besar dan stabil dibandingkan algoritme lainnya. Akan tetapi, dalam beberapa kondisi, algoritme NBC menghasilkan nilai akurasi yang rendah apabila fitur atau parameter data ditambahkan [3]. Selain itu, algoritme *naive Bayes* dapat bekerja dengan baik dalam mengenali aktivitas fisik manusia secara *realtime*. Disebutkan juga bahwa NBC lebih unggul dalam waktu pemrosesan dibandingkan algoritme kecerdasan buatan populer lainnya, seperti *Support Vector Machine* (SVM), *OneR*, *neural network*, dan *decision tree* [4].

Penelitian yang telah dilakukan sebelumnya menunjukkan bahwa algoritme NBC mempunyai kinerja yang baik dalam melakukan klasifikasi serta mudah diterapkan. *Naive Bayes* merupakan algoritme *supervised learning*, sehingga dalam klasifikasi data membutuhkan data latih yang cukup untuk mendapatkan hasil yang baik [5].

Fokus penelitian ini adalah penentuan klasifikasi aktivitas berkendara yang meliputi berhenti, lurus, belok kanan, dan belok kiri, menggunakan algoritme klasifikasi *naive Bayes*. Algoritme NBC bekerja dengan melakukan klasifikasi data dan memprediksi peluang data uji yang akan terjadi berdasarkan data latih sebelumnya [6]. Data diperoleh berdasarkan

<sup>1</sup> Teknik Telekomunikasi Politeknik Elektronika Negeri Surabaya, Jl. Raya ITS Politeknik Elektronika, Sukolilo, Keputih, Surabaya Jawa Timur, INDONESIA 60111 (telp: (031) 5947280; fax: 031-5946114; e-mail: haniah @pens.ac.id)



Gbr. 1 Ilustrasi cara kerja akselerometer pada *smartphone*.



Gbr. 2 Ilustrasi sensor giroskop pada *smartphone*.

perubahan sumbu  $x$ ,  $y$  dan  $z$  dari akselerometer dan giroskop yang mengidentifikasi kemiringan kendaraan. Algoritme NBC bekerja dengan menghitung nilai probabilitas setiap kelas dan menggolongkan secara linier dengan sangat efisien. Kinerja algoritme NBC relatif kuat, mudah diterapkan, cepat, dan akurat [7]. Metode *smoothing* berfungsi untuk meningkatkan klasifikasi pada NBC agar nilai akurasi yang dihasilkan lebih besar [8].

Penjabaran secara detail makalah ini dibagi menjadi beberapa bagian. Bagian pertama menjabarkan pendahuluan dan latar belakang, bagian kedua hingga keenam menjabarkan teori yang terkait dengan penelitian, bagian ketujuh menjabarkan metodologi dan cara kerja sistem, bagian kedelapan menjabarkan hasil penelitian, serta bagian kesembilan berisi kesimpulan.

## II. SENSOR AKSELEROMETER *SMARTPHONE*

Sensor akselerometer bekerja dengan mengukur nilai kecepatan pada gerakan penggunaannya dalam tiga arah (sumbu  $x$ ,  $y$ , dan  $z$ ) yang dapat diproses lebih lanjut untuk mengenali pola aktivitas pengendara [9]. Sensor akselerometer juga dapat mengukur nilai kecepatan, getaran atau vibrasi, serta percepatan akibat gravitasi. Sensor akselerometer akurat untuk mengukur kemiringan sudut sistem yang diam atau tidak bergerak, tetapi tidak akurat untuk mengukur kemiringan sudut sistem yang bergerak. Data kemiringan sudut memiliki derau. Hal ini terjadi karena untuk mendapatkan data kemiringan sudut dari akselerometer, dibutuhkan data percepatan gravitasi bumi [9], tetapi jika sistem bergerak, akselerometer juga mengukur percepatan gerak sistem tersebut, sehingga pengukuran tidak akurat. Sensor akselerometer dapat mengukur percepatan berdasarkan perubahan koordinat sumbu, seperti diilustrasikan pada Gbr. 1.

Pengujian ini memanfaatkan jenis sensor tipe *linear acceleration* untuk pengambilan *dataset* serta pengujiannya. Sensor tipe *linear acceleration* bekerja dengan tidak dipengaruhi gravitasi bumi, sehingga dapat membaca tiga sumbu layaknya sensor giroskop. Secara teori, akselerometer bekerja seperti halnya sebuah sistem massa pegas sederhana yang sesuai dengan hukum Hooke seperti pada (1).

$$F = k \cdot x \tag{1}$$

dengan  $F$  merupakan gaya pegas,  $k$  adalah nilai konstanta pegas, dan  $x$  adalah simpangan pada pegas. Dari persamaan

hukum Hooke tersebut, diperoleh persamaan hukum kedua Newton seperti pada (2).

$$F = m \cdot a \tag{2}$$

dengan  $m$  merupakan massa benda dalam satuan kg, sedangkan  $a$  adalah nilai percepatan dalam  $m/s^2$ . Jika sistem ini mengalami percepatan, maka sesuai dengan hukum Newton, akan timbul gaya yang menggantikan massa. Gaya ini menyebabkan massa bergeser. Oleh karena itu, percepatan akan menyebabkan massa menjadi tergeser, sesuai dengan (3).

$$x = \frac{m \cdot a}{k} \tag{3}$$

Dengan mengamati perpindahan  $x$ , diketahui bahwa massa mengalami percepatan  $a$ , seperti pada (4).

$$a = \frac{k \cdot x}{m} \tag{4}$$

Secara konseptual, sensor akselerometer menentukan nilai percepatan yang diterapkan pada perangkat ( $AD$ ) dengan mengukur gaya yang terdapat pada sensor itu sendiri ( $F_s$ ), menggunakan (5).

$$AD = -\left(\frac{1}{\text{massa}}\right) \sum F_s \tag{5}$$

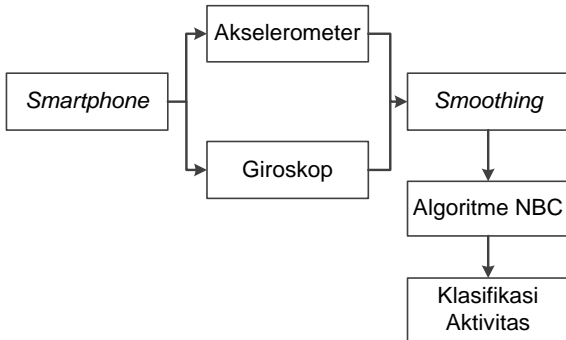
## III. SENSOR GIROSKOP *SMARTPHONE*

Sensor giroskop merupakan sensor yang berfungsi mengukur rotasi atau perputaran suatu perangkat berdasarkan gerakan. Giroskop bekerja dengan mengukur laju rotasi dalam satuan ( $rad/s$ ) di sekitar sumbu  $x$ ,  $y$ , dan  $z$  *smartphone*, seperti tampak pada Gbr. 2. Sensor giroskop akurat untuk mengukur kemiringan sudut sistem yang bergerak atau pada kondisi dinamik, karena keluaran giroskop tidak dipengaruhi oleh percepatan translasi dan getaran. Namun, dari sifat alami giroskop yang mempunyai suatu *bias error*, apabila integral terhadap waktu dilakukan, kemungkinan akan terjadi akumulasi dari *error* tersebut, yang mengakibatkan hasil pengukuran posisi sudut giroskop cenderung melayang (*drifting*) dari waktu ke waktu. Oleh karenanya, giroskop bekerja dengan baik apabila terdapat sensor akselerometer [10].

Sensor giroskop yang terdapat pada *smartphone* berbasis *Micro Electro Mechanical Systems* (MEMS) yang tertanam dalam *chip* atau *Integrated Circuit* (IC). Sensor ini berukuran kecil, sehingga mudah digunakan atau diaplikasikan (*wearable*) dan memiliki konsumsi daya yang rendah [10].



Gbr. 3 Proses perhitungan algoritme NBC.



Gbr. 4 Perancangan sistem.



Gbr. 5 Posisi smartphone.

IV. K-FOLD CROSS VALIDATION

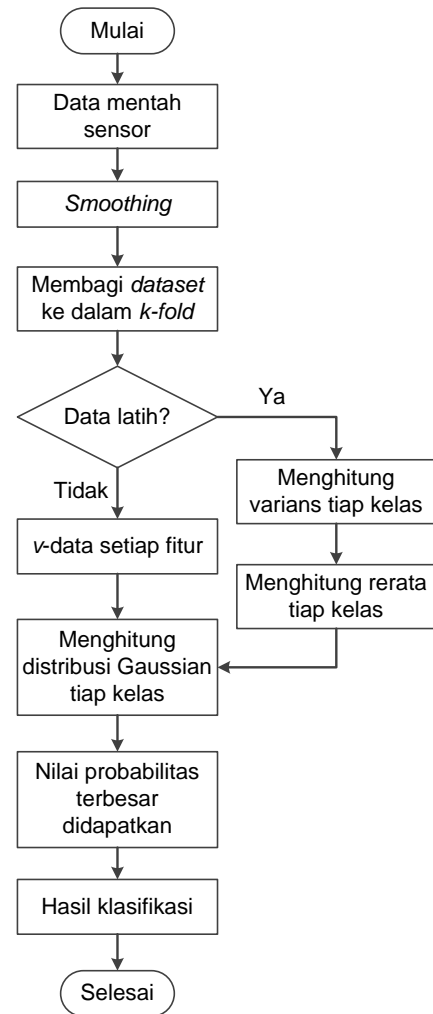
Cross validation merupakan salah satu teknik dalam menilai keakuratan suatu dataset baru guna membangun klasifikasi atau model prediksi untuk menilai suatu model, yakni data latih dan data uji. Data latih merupakan kumpulan suatu data yang diperoleh dari hasil pengukuran yang disebut dengan data mentah (raw), kemudian dilakukan filter hingga menjadi dataset atau data bersih. Data latih digunakan dalam proses pembangunan model, sedangkan data uji digunakan untuk validasi model tersebut.

Salah satu teknik yang dapat digunakan adalah k-fold cross validation. Dalam teknik k-fold cross validation, dataset dibagi sesuai jumlah k buah partisi secara acak. Kemudian, dilakukan eksperimen sejumlah k kali, dengan setiap eksperimen menggunakan sejumlah data ke-k sebagai data uji dan sisa data lainnya digunakan sebagai data latih [11].

Metode fold cross validation digunakan untuk menilai kinerja model dengan menghitung tingkat kesalahan prediksi pada model. Selain itu, k-fold juga berfungsi agar tidak terdapat overlapping data terhadap data uji. Validasi pada makalah ini menggunakan 2 k-fold hingga 10 k-fold, dengan akurasi yang dihasilkan setiap k-fold dibandingkan guna mendapatkan nilai akurasi yang terbaik.

V. SINGLE EXPONENTIAL SMOOTHING

Salah satu metode peramalan yang sering digunakan adalah exponential smoothing. Metode smoothing merupakan sebuah



Gbr. 6 Implementasi algoritme NBC.

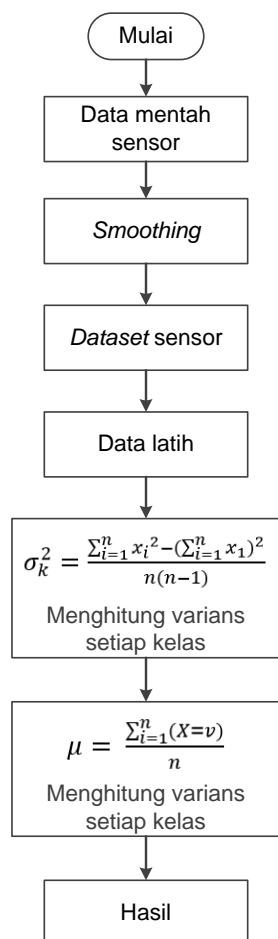
teknik yang dapat memisahkan sinyal dan derau, dalam hal ini sinyal mewakili pola pada data, sedangkan derau merupakan kesalahan pada waktu ke-t. Dengan demikian, Single Exponential Smoothing (SES) akan bekerja secara kontinu memperbaiki prediksi data dengan menghitung rata-rata nilai dari data sebelumnya pada suatu data berupa deret waktu dengan cara menurun (eksponensial). Metode SES tepat digunakan untuk peramalan data yang mengalami fluktuasi secara acak (tidak teratur) [12]. Perhitungan SES dilakukan dengan menghitung selisih angka peramalan dengan angka sebelumnya, seperti pada (6).

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) \tag{6}$$

dengan  $F_t$  merupakan hasil peramalan,  $F_{t-1}$  merupakan hasil peramalan sebelumnya,  $A_{t-1}$  adalah data aktual sebelumnya, sedangkan  $\alpha$  merupakan konstanta smoothing sebagai parameter model dalam mengurangi faktor acak suatu data.

VI. NAÏVE BAYES CLASSIFIER

Naïve Bayes merupakan sebuah algoritme yang mengklasifikasi suatu data atau nilai dengan menghitung kelompok probabilitas dan kemudian menjumlahkan



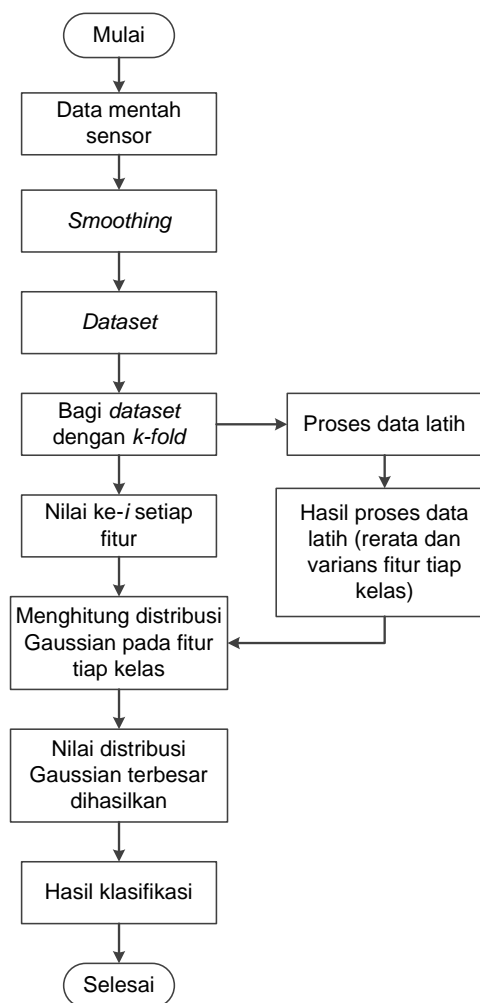
Gbr. 7 Proses pelatihan data algoritme *naïve Bayes*.

kombinasi nilai dari *dataset* yang telah dikumpulkan. Aplikasi teorema Bayes banyak ditemukan pada bidang komputer cerdas sebagai salah satu dasar dari metode *machine learning* dan *data mining*. Dalam kinerjanya, *naïve Bayes* tidak membutuhkan jumlah data latih yang besar dalam menentukan estimasi parameter yang diperlukan pada proses klasifikasi. Hal ini merupakan salah satu keuntungan algoritme NBC. *Naïve Bayes* bekerja dengan cukup baik dalam kebanyakan persoalan di dunia nyata yang kompleks daripada yang diharapkan. Pada beberapa penelitian, penerapan algoritme *naïve Bayes* menggunakan perhitungan distribusi untuk *event model*, sesuai kebutuhan analisis data yang digunakan. Apabila *dataset* yang digunakan bersifat kontinu dan terkait antar kelasnya, maka digunakan rumus perhitungan distribusi normal atau yang biasa disebut sebagai distribusi Gaussian [13]. Distribusi Gaussian merupakan distribusi dari variabel acak kontinu. Distribusi ini paling banyak dan paling sering digunakan.

Proses perhitungan algoritme NBC ditunjukkan pada Gbr. 3. Penyelesaian perhitungan algoritme NBC dilakukan sebagai berikut.

A. Menghitung Likelihood

*Likelihood* merupakan nilai peluang terjadinya karakteristik-karakteristik sampel pada kelas. Dengan kata lain, *likelihood* merupakan nilai kemungkinan yang dipertanyakan.



Gbr. 8 Proses pengujian data algoritme *naïve Bayes*.

Perhitungan *likelihood* dilakukan secara independen untuk setiap atribut. *Likelihood* untuk data latih dapat dihitung dengan mencari *rata-rata* dan varians masing masing atribut. Untuk menghitung nilai *likelihood*, dibutuhkan nilai rata-rata ( $\mu$ ) dan varian ( $\sigma^2$ ) dari data latih, seperti pada (7).

$$\mu = \frac{\sum_{i=1}^n (X=v)}{n} \tag{7}$$

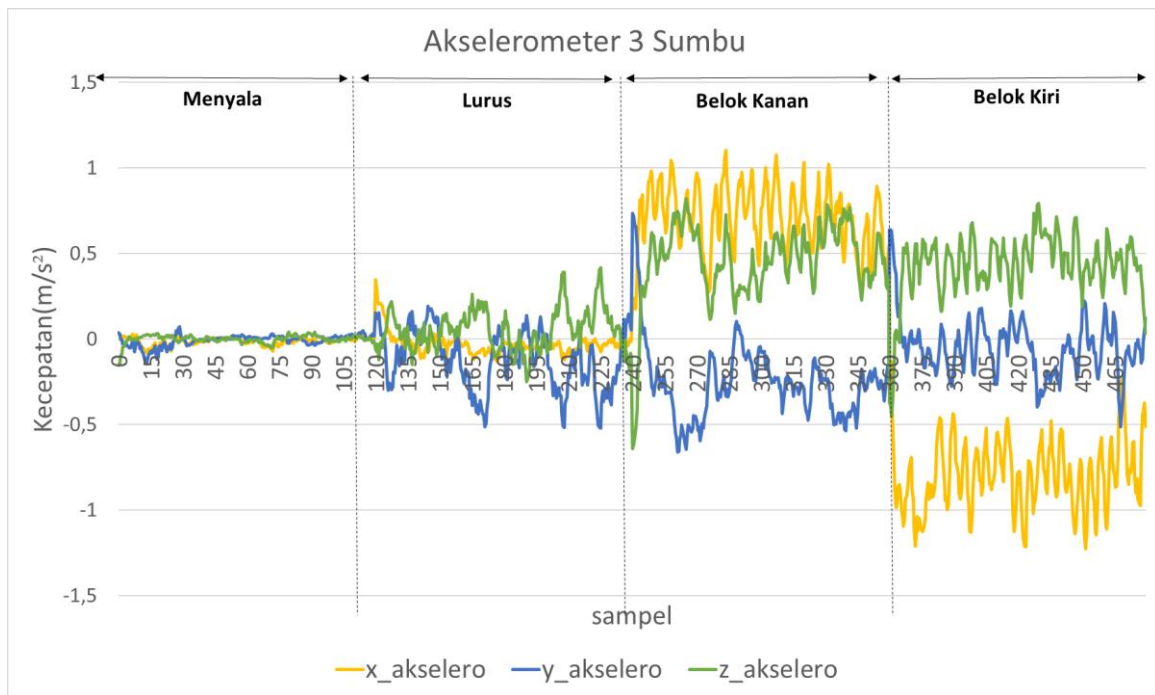
Untuk menentukan varians ( $\sigma_k^2$ ), digunakan (8).

$$\sigma_k^2 = \frac{\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)} \tag{8}$$

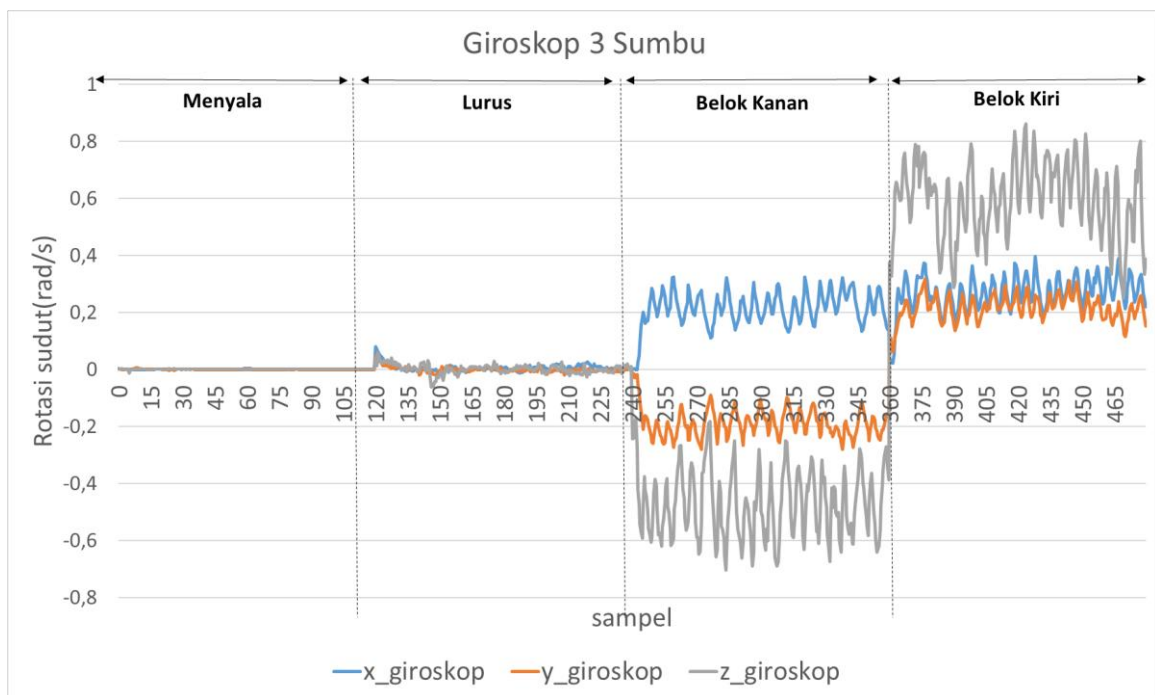
dengan  $\mu$  merupakan rata-rata sampel,  $\sigma_k^2$  adalah varian,  $n$  adalah ukuran sampel, dan  $x_1$  merupakan data ke- $i$ .

B. Distribusi Gaussian

Klasifikasi *naïve Bayes* dilakukan menggunakan distribusi Gaussian. Pada perhitungan distribusi ini dibutuhkan dua parameter, yaitu  $\mu$  dan  $\sigma_k^2$ . Fungsi densitas menunjukkan nilai probabilitas relatif. Pada distribusi Gaussian, diasumsikan jika nilai kontinu terkait dengan setiap kelas, maka didistribusikan menggunakan distribusi normal (Gaussian). Untuk menghitung nilai *likelihood* data uji kontinu, digunakan (9).



(a)



(b)

Gbr. 9 Grafik sensor *smartphone* dalam domain waktu, (a) akselerometer, (b) giroskop.

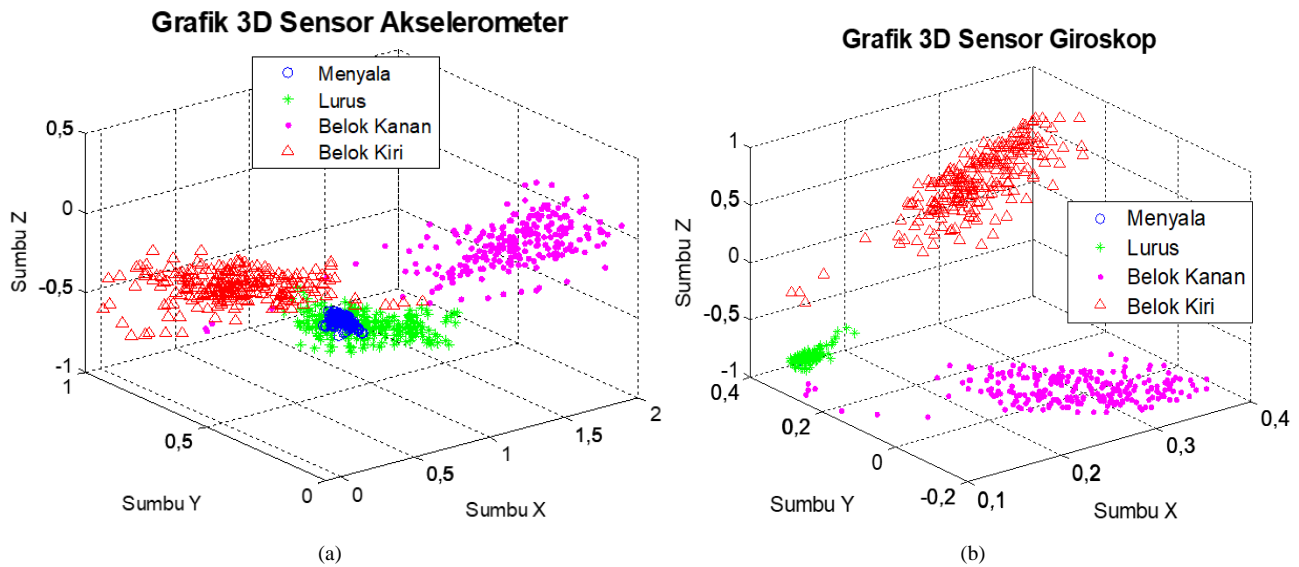
$$P(X = v|C) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(v-\mu_k)^2}{2\sigma_k^2}} \quad (9)$$

dengan  $V$  merupakan nilai fitur ke- $i$  pada data baru,  $\sigma_k^2$  adalah varian data setiap fitur untuk tiap kelas, dan  $\mu_k$  adalah nilai rata-rata untuk tiap fitur pada masing-masing kelas.

## VII. METODOLOGI

Proses perancangan dan pembuatan sistem ditunjukkan pada Gbr. 4. Perubahan sensor akselerometer dan giroskop diperoleh melalui pengambilan data menggunakan *smartphone*. *Dataset* berupa data sensor akselerometer dan giroskop yang terkumpul dan disimpan dalam file *.csv*. Pada *dataset* kemudian dilakukan





Gbr. 10 Grafik 3D sensor *smartphone*, (a) akselerometer, (b) giroskop.

TABEL I  
AKURASI DATA MENGGUNAKAN *K-FOLD*

<i>K-fold</i>	NBC		
	Data Mentah	SES $\alpha = 0,1$	SES $\alpha = 0,9$
2	98,12%	98,54%	98,12%
3	98,33%	98,54%	98,33%
4	98,33%	98,95%	98,54%
5	98,22%	99,16%	98,33%
6	98,43%	99,16%	98,33%
7	98,43%	99,27%	98,33%
8	98,43%	99,27%	98,43%
9	98,32%	99,26%	98,32%
10	98,43%	99,16%	98,33%

TABEL II  
HASIL AKURASI DATA

Dataset	Akurasi
Sensor giroskop dan akselerometer tanpa <i>smoothing</i>	98,43%
Sensor giroskop dan akselerometer <i>smoothing</i> $\alpha = 0,1$	99,27%
Sensor giroskop dan akselerometer <i>smoothing</i> $\alpha = 0,9$	98,43%

*smoothing* sebelum proses pelatihan data dengan Python untuk mendapatkan klasifikasi aktivitas berkendara. Pengumpulan data dilakukan untuk mendapatkan klasifikasi aktivitas pengendara seperti belok kanan, belok kiri, lurus, serta mesin menyala. Proses pengambilan *dataset* dilakukan dengan meletakkan *smartphone* sejajar dengan *dashboard* sepeda motor, seperti yang terlihat pada Gbr. 5.

Pengumpulan data ini dikategorikan ke dalam data latih, yakni data yang akan dilatih pada algoritme, sehingga dapat menentukan kelas data uji. *Dataset* yang diambil berupa data sensor akselerometer dan giroskop *smartphone*. Proses *smoothing* dilakukan sebelum data melewati proses klasifikasi. *Dataset* sangat fluktuatif, sehingga dibutuhkan proses *smoothing*, yang dapat mereduksi *dataset* sehingga data yang dihasilkan lebih stabil dan terstruktur, sebelum dilakukan klasifikasi. Proses *smoothing* dilakukan menggunakan SES

dengan parameter nilai  $\alpha$  maksimum dan minimum, yakni 0,1 dan 0,9. Proses klasifikasi aktivitas berkendara menggunakan klasifikasi NBC diperlihatkan pada Gbr. 6.

Pada proses NBC terdapat dua tahapan, yaitu proses pelatihan data dan proses pengujian data. Kedua proses dijelaskan sebagai berikut.

1) *Proses Pelatihan Data*: Proses pelatihan data dengan menggunakan algoritme NBC ini dipengaruhi oleh *likelihood* dan *prior*. Proses pelatihan data ini melewati proses *k-fold* sebagai bentuk evaluasi hasil pelatihan, kemudian dihitung menggunakan metode *k-fold cross validation*. Pada metode ini, data latih dievaluasi sejumlah *k-subset* yang dibentuk. Cara metode *k-fold* bekerja yakni dengan membagi data sejumlah *k*, kemudian melakukan *iterasi* data uji pada data latih sebanyak *k subset* juga.

Pada proses pelatihan data, data *input* hasil *smoothing* dihitung menggunakan persamaan distribusi Gaussian. Nilai rata-rata dan varians dihitung berdasarkan sejumlah kelas untuk setiap fitur. Perhitungan algoritme NBC berhenti apabila telah ditemukan nilai probabilitas yang paling besar dari semua kelas. Proses pelatihan data ditunjukkan pada Gbr. 7.

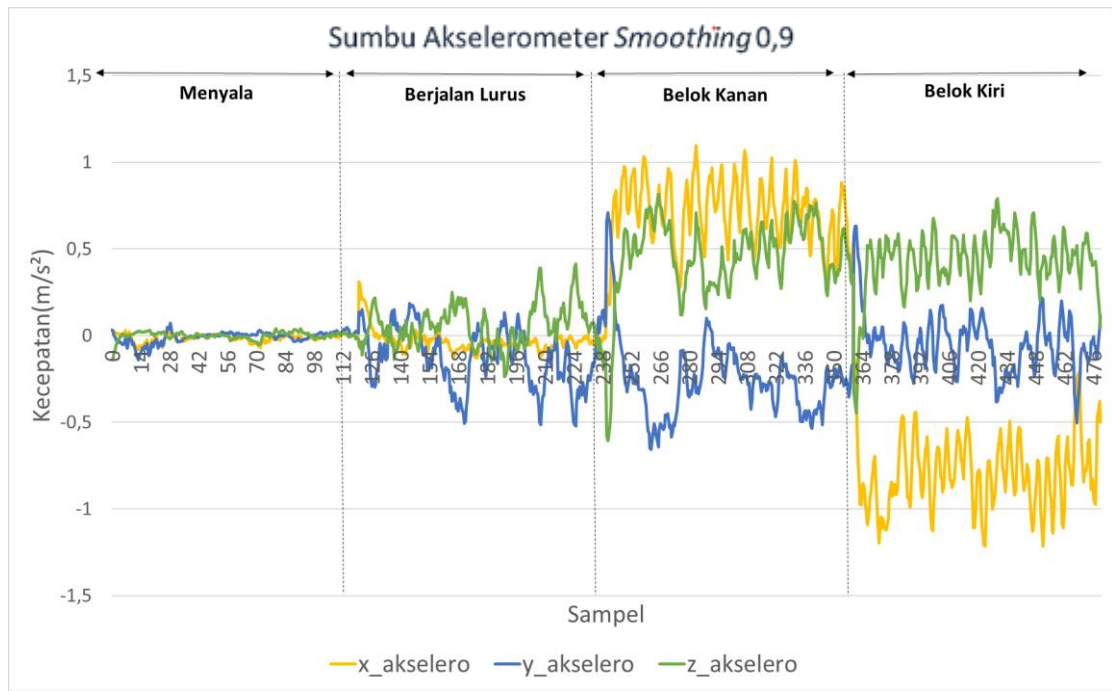
2) *Proses Pengujian Data*: Proses pengujian data dilakukan dengan mencari nilai probabilitas terbesar dari setiap kelas setelah dihasilkan nilai variasi dan rata-rata setiap kelas dari proses pelatihan, dengan data yang dihasilkan digunakan untuk proses. Gbr. 8 merupakan proses pengujian yang dilakukan.

### VIII. HASIL PENELITIAN

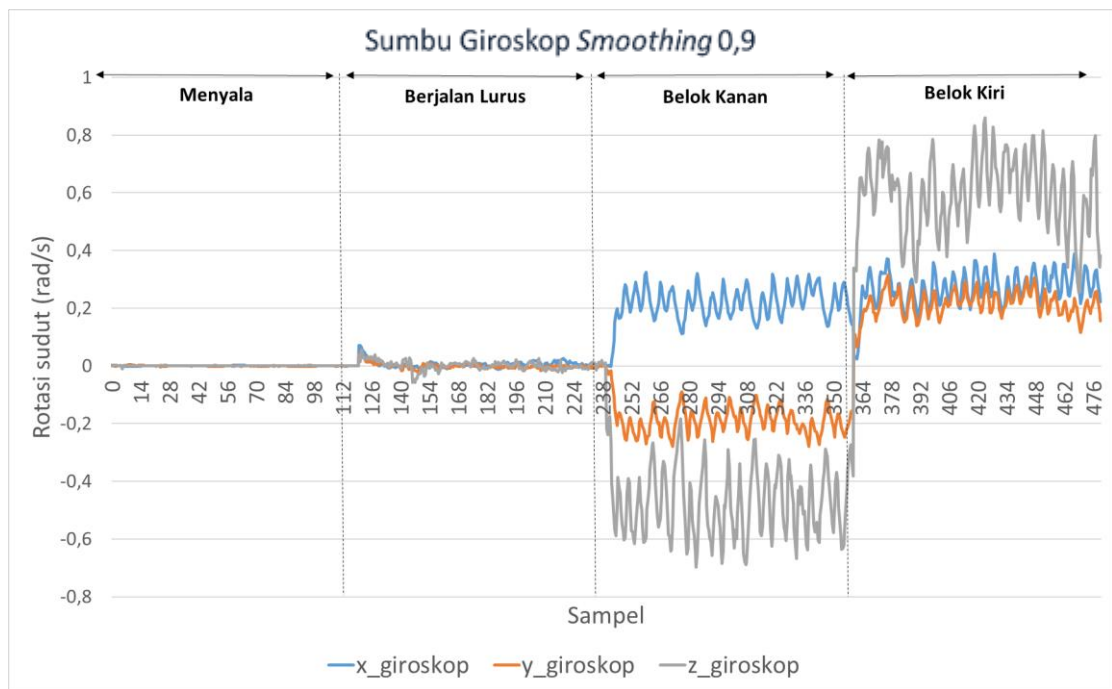
Pengujian dilakukan untuk mengidentifikasi aktivitas berkendara menggunakan algoritme NBC dengan memanfaatkan sensor akselerometer dan giroskop pada *smartphone*.

#### A. Pengujian Dataset

*Dataset* diperoleh dari data sensor akselerometer dan giroskop *smartphone* pada domain waktu, seperti terlihat pada



(a)

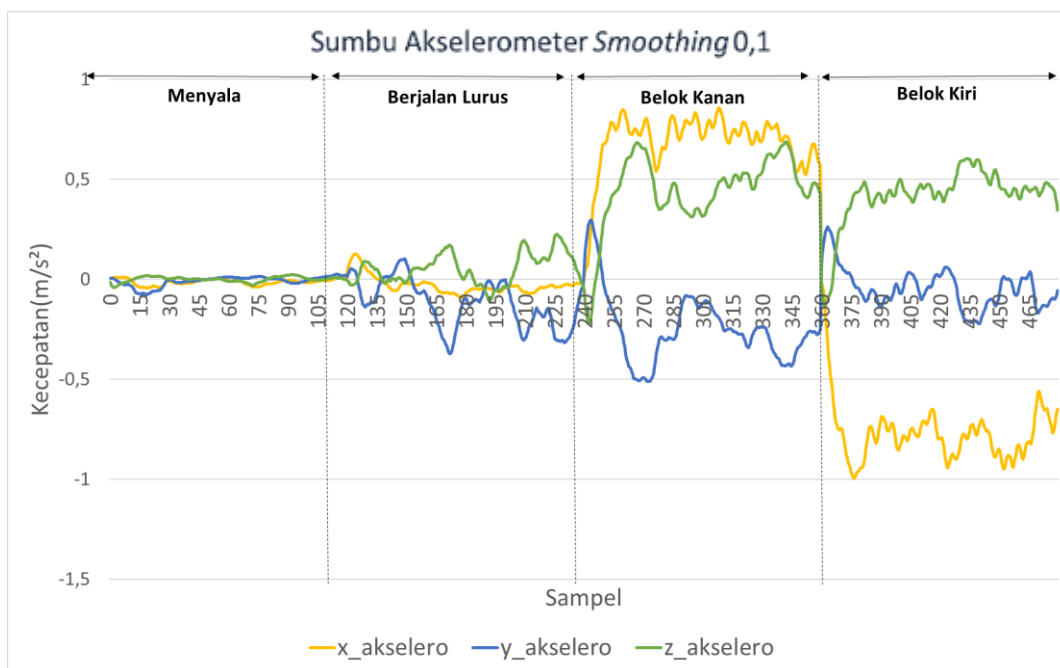


(b)

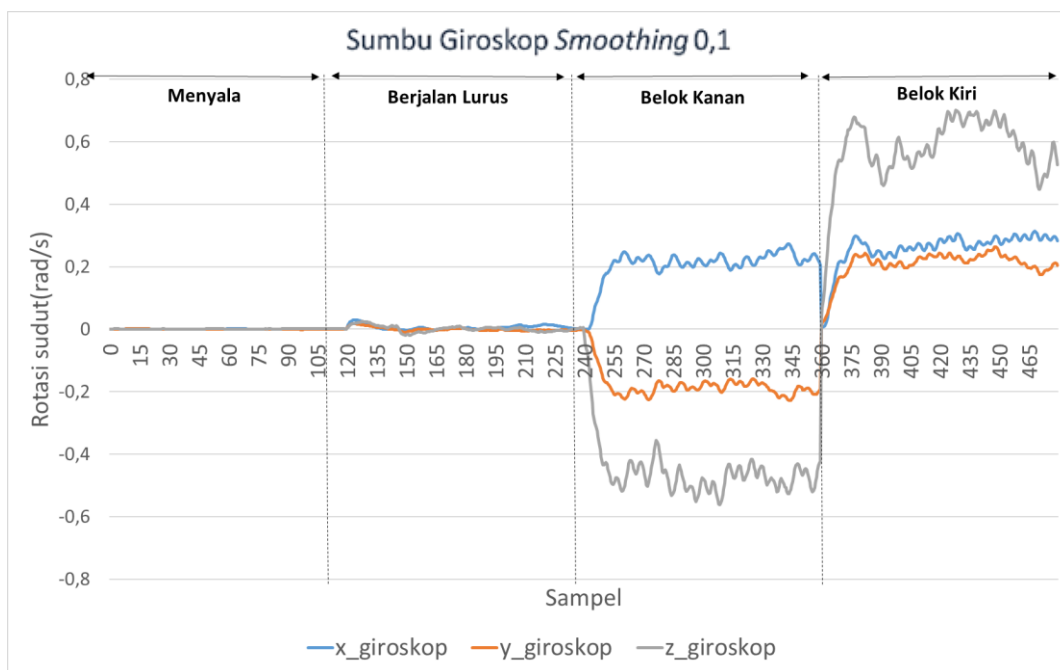
Gbr. 11 Grafik sensor *smartphone* setelah proses *smoothing* dengan  $\alpha = 0,9$ , (a) akselerometer, (b) giroskop.

Gbr. 9 dan Gbr. 10. Dari gambar tersebut diketahui bahwa *dataset* yang diperoleh mempunyai data yang fluktuatif. Oleh karena itu, diperlukan *smoothing* untuk mereduksi data mentah, sehingga data lebih stabil. Setiap aktivitas memiliki karakteristik yang berbeda satu sama lain. Dalam proses ini, digunakan nilai  $\alpha$  maksimum dan minimum, yakni 0,9 dan 0,1. Hasil reduksi menggunakan SES diperlihatkan pada Gbr. 11 dan Gbr. 12.

Terlihat dari Gbr. 11 dan Gbr. 12 bahwa nilai  $\alpha$  yang digunakan memengaruhi hasil peramalan. Nilai  $\alpha$  yang besar menunjukkan tingkat kepercayaan data sebelum periode peramalan besar. Demikian pula sebaliknya, jika nilai  $\alpha$  kecil, maka tingkat kepercayaan data sebelum periode peramalan kecil. Oleh karena itu, *smoothing* menggunakan  $\alpha = 0,1$  menghasilkan lebih banyak selisih data hasil peramalan dibandingkan menggunakan  $\alpha = 0,9$ .



(a)



(b)

Gbr. 12 Grafik sensor *smartphone* setelah proses *smoothing* dengan  $\alpha = 0,1$ . (a) akselerometer (b) giroskop.

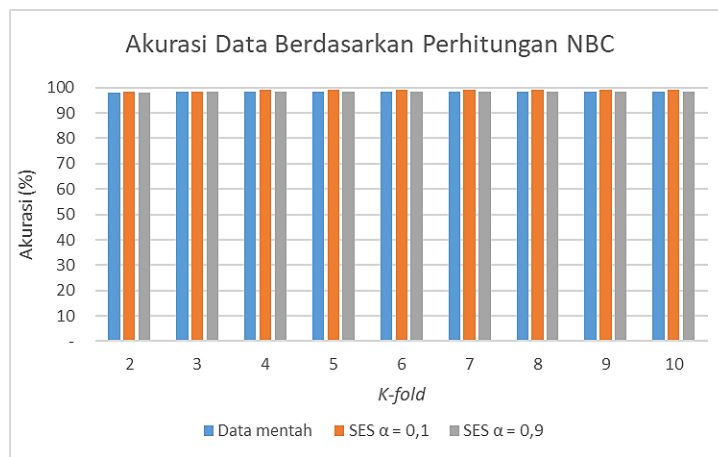
### B. Klasifikasi Aktivitas Berkendara

Proses klasifikasi dilakukan dengan dua parameter, kemudian hasilnya dibandingkan, yakni yang pertama melalui proses *smoothing*, sebelum dilakukan klasifikasi menggunakan algoritme NBC, dan yang kedua klasifikasi NBC tanpa proses *smoothing* terlebih dahulu. Setiap aktivitas mengemudi memiliki pola yang berbeda, baik pada sensor akselerometer dan giroskop. Pola ini kemudian digunakan sebagai data

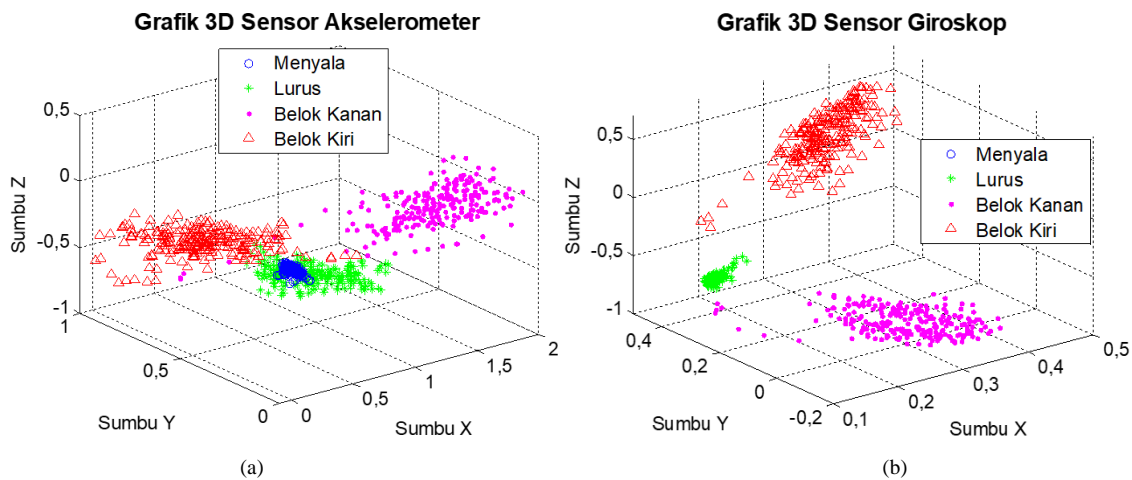
referensi untuk NBC dalam menentukan kelas klasifikasi aktivitas mengemudi. Pada Tabel I disajikan hasil akurasi pengujian data sensor *smartphone* menggunakan *smoothing* dan tanpa *smoothing* yang dihasilkan dari perbandingan data uji dengan data latih 2:8, yakni setiap dua data uji dihitung berdasarkan referensi sebanyak delapan data latih.

Dari Tabel I diketahui nilai akurasi aktivitas berkendara yang dihasilkan dari proses klasifikasi NBC dengan beberapa nilai *k-fold*. Nilai *k-fold* yang digunakan adalah antara 2 sampai

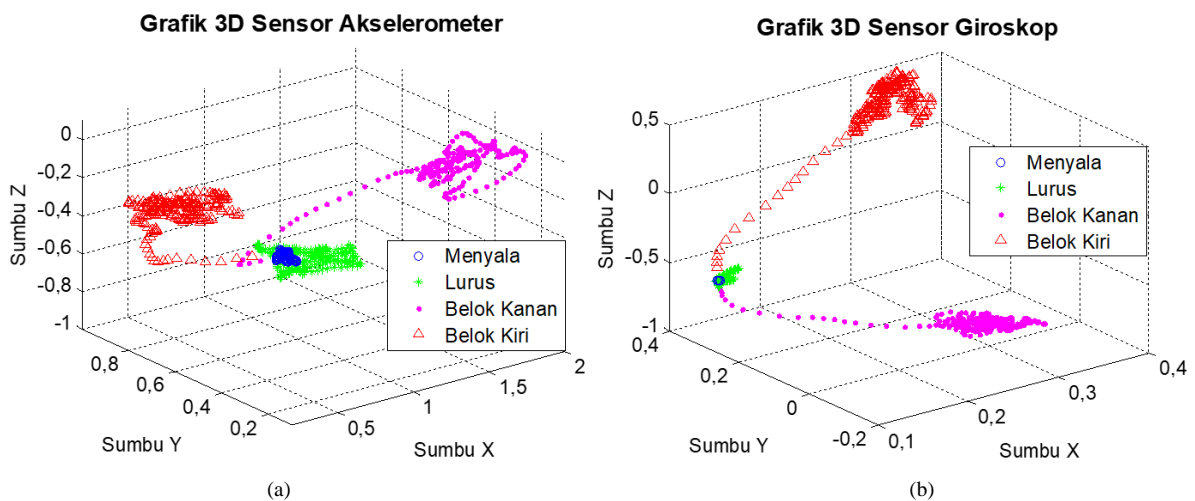




Gbr. 13 Hasil perhitungan NBC tanpa proses *smoothing* dan dengan proses *smoothing* dengan nilai  $\alpha = 0,1$  dan  $0,9$ .



Gbr. 14 Grafik sensor *smartphone* 3D setelah proses *smoothing* dengan  $\alpha = 0,9$ , (a) akselerometer, (b) giroskop.



Gbr. 15 Grafik sensor *smartphone* 3D setelah proses *smoothing* dengan  $\alpha = 0,1$ , (a) akselerometer, (b) giroskop.

10. Proses NBC mempunyai nilai akurasi yang tinggi pada  $k$ -fold = 8. Perhitungan menggunakan data mentah tanpa dan melalui proses *smoothing*  $\alpha = 0,9$  menghasilkan akurasi yang sama, yakni 98,43%. Jika dikaitkan dengan grafik yang telah ditunjukkan pada Gbr. 13, terlihat bahwa tidak banyak nilai yang direduksi dan cenderung menyerupai data mentah

sebelum proses *smoothing*. Semakin besar nilai  $\alpha$  yang digunakan, semakin sedikit nilai yang direduksi serta data yang dihasilkan lebih fluktuatif. Sedangkan akurasi yang tertinggi aktivitas berkendara melalui proses *smoothing* dengan  $\alpha = 0,1$  adalah 99,27%, seperti terlihat pada Tabel II. Hal ini berbanding lurus dengan banyak nilai yang direduksi pada saat

proses *smoothing*, dan data yang dihasilkan setelah proses *smoothing*  $\alpha = 0,1$  lebih stabil. Grafik 3D dimensi posisi sensor akselerometer dan giroskop dengan sumbu x, y, dan z ditunjukkan pada Gbr. 14 dan Gbr. 15.

#### IX. KESIMPULAN

NBC berhasil melakukan klasifikasi identifikasi aktivitas berkendara dengan memanfaatkan sensor akselerometer dan giroskop pada *smartphone*. Proses klasifikasi aktivitas berkendara menggunakan algoritme NBC yang dikombinasikan dengan metode SES menghasilkan nilai akurasi yang lebih tinggi. Nilai akurasi yang tertinggi pada data sensor akselerometer dan giroskop dengan *smoothing*  $\alpha = 0,1$  adalah 99,27%, dengan menggunakan *k-fold* = 8, yaitu perbandingan setiap dua data uji menggunakan delapan referensi data latihan.

#### UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Politeknik Elektronika Negeri Surabaya yang telah memberikan semangat dan dukungan pada penelitian serta seluruh dosen anggota Lab JJ-305 yang telah memberikan motivasi dan semangat.

#### REFERENSI

- [1] Social Determinants of Health, *Global Status Report on Road Safety 2018*, Geneva, Switzerland: World Health Organization, 2018.
- [2] (2020) "Rata-rata Tiga Orang Meninggal Setiap Jam Akibat Kecelakaan Jalan," [Online], [https://kominfo.go.id/index.php/content/detail/10368/rata-rata-tiga-orang-meninggal-setiap-jam-akibat-kecelakaan-jalan/0/artikel\\_gpr](https://kominfo.go.id/index.php/content/detail/10368/rata-rata-tiga-orang-meninggal-setiap-jam-akibat-kecelakaan-jalan/0/artikel_gpr), tanggal akses: 1-Jan-2020.
- [3] L. Dang-Nhae, N. Due-Nhan, N. Thi-Hau, dan N. Ha-Nam, "Vehicle Mode and Driving Activity Detection Based on Analyzing Sensor Data of Smartphone," *Sensors*, Vol. 18, No. 4, hal. 1-25, Mar. 2018.
- [4] A. Dinh, L. Chen, dan X. Yang, "Implementation of a Wearable Real-Time System for Physical Activity Recognition based on Naive Bayes Classifier," *IEEE International Conference on Bioinformatics and Biomedical Technology*, 2010, hal. 101-105.
- [5] J. Santoso, A.D.B. Soetiono, Gunawan, E. Setyati, E.M. Yuniarno, M. Hariadi, dan M.H. Purnomo, "Self-Training Naive Bayes Berbasis Word2Vec untuk Kategorisasi Berita Bahasa Indonesia," *JNTEFI*, Vol. 7, No. 2, hal. 158-166, Mei 2018.
- [6] H. Zhang dan D. Li, "Naive Bayes Text Classifier," *2007 IEEE International Conference on Granular Computing (GRC 2007)*, 2007, hal. 708-711.
- [7] F.-J. Yang, "An Implementation of Naive Bayes Classifier," *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018, hal. 301-306.
- [8] Q. Yuan, G. Cong, dan N.M. Thalmann, "Enhancing Naive Bayes with Various Smoothing Methods for Short Text Classification," *Proceedings of the 21st International Conference on World Wide Web*, 2012, hal. 645-646.
- [9] G. Dewantoro, F.D. Setiaji, dan E.C. Mone, "Skema Kendali Ball-Balancing Robot Secara Nirkabel," *JNTEFI*, Vol. 3, No. 2, hal. 136-141, Mei 2014.
- [10] R. Setiawan, A. Arifin, F. Budiman, dan Ad. Soeprijanto, "Desain Sistem Pengukuran Lower Limb Joint Angles pada Kondisi Dinamik untuk FES," *JNTEFI*, Vol. 7, No. 1, hal. 112-121, Feb. 2018.
- [11] Y. Wihardi (2013) "K-Folds Cross Validation" [Online], <http://blog.yayaw.web.id/riset/k-folds-cross-validation>, tanggal akses: 6-Mei-2020.
- [12] R. Angrainingsih, G.R. Aprianto, dan S.W. Sihwi, "Time Series Forecasting Using Exponential Smoothing to Predict the Number of Website Visitor of Sebelas Maret University," *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2015, hal. 14-19.
- [13] S. Raschka, "Naive Bayes and Text Classification I: Introduction and Theory," *eprint arXiv:1410.5329*, hal. 1-20, 2014.