



Traffic Intensity Detection in Lagos State Using Bayesian Estimation Model

Aaron A. Izang^{1*}, Seun O. Arowosegbe², Oluwabukola F. Ajayi², Aderonke Adegbenjo²,
Onome B. Ohwo², Afolarin I. Amusa², Wumi S. Ajayi³, Alfred A. Udosen²

¹ Department of Information Technology, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo 121103, Nigeria

² Department of Computer Science, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo 121103, Nigeria

³ Department of Software Engineering, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo 121103, Nigeria

Corresponding Author Email: izanga@babcock.edu.ng

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ijcmem.120412>

ABSTRACT

Received: 21 November 2023

Revised: 21 May 2024

Accepted: 31 July 2024

Available online: 27 December 2024

Keywords:

Bayesian model, Gaussian estimation, management, multinomial, traffic intensity

Traffic congestion is a significant challenge in Lagos State, Nigeria. Existing methods rely on limited data sources and simplistic models that fail to capture the complexities of traffic dynamics in a congested urban environment. This study focused on traffic intensity detection in Lagos State using a Bayesian estimation model. Data was obtained from the Kaggle website which involved observing the number of vehicles intersecting junctions at various times of the day for a week. The model captured both spatial and temporal variations, providing real-time estimations of traffic congestion levels across different road segments. Comparative analysis with existing traffic estimation methods showed superior performance in terms of accuracy and reliability. The Gaussian Naive Bayes model achieved a high accuracy of 96% and balanced f1-score of 96%, precision of 0.96, and recall of approximately 0.96. On the other hand, the multinomial Naive Bayes model achieved an accuracy of 69% with a lower f1-score of 69%, precision of 0.67, and recall of 0.69. The model's capacity to provide accurate real-time site traffic facts can significantly contribute to effective traffic control and concrete making plans initiatives.

1. INTRODUCTION

Traffic congestion is a pervasive trouble in city areas, and Lagos State, Nigeria, is no exception. The rapidly growing populace and growing car possession costs have caused extreme site traffic congestion in Lagos State, causing large economic losses, environmental pollution, and decreased great of lifestyles for residents [1]. Accurate real-time traffic facts perform a critical role in effective traffic management and concrete making plans. However, the contemporary techniques for detecting and predicting traffic intensity in Lagos State lack accuracy and reliability, hindering the implementation of focused congestion mitigation techniques [2, 3]. Therefore, this study aims to develop a Bayesian estimation model to accurately identify traffic intensity levels in Lagos State.

The lack of correct and well-timed site traffic records poses a widespread task in dealing with traffic congestion in Lagos State. Existing techniques depend on limited facts assets and simplistic models that fail to seize the complexities of site traffic dynamics in congested urban surroundings. As an end result, site traffic control authorities war to make informed selections and put in force proactive measures to alleviate congestion [4]. There is a critical want for an advanced

estimation model that leverages comprehensive information assets and incorporates superior statistical techniques to offer correct and actual-time site traffic depth records for effective congestion management [5].

Existing research inside the discipline of traffic depth detection have predominantly focused on advanced towns with properly-established site traffic control systems. Limited studies have been carried out specifically for Lagos State, which possesses specific traffic patterns and demanding situations. However, research making use of Bayesian estimation fashions have tested promising effects in accurately estimating site traffic intensity levels in different urban regions.

The improvement and implementation of a correct traffic intensity detection version in Lagos State could have sizeable implications for site traffic management and concrete planning. Accurate real-time traffic records will enable traffic government to make informed choices, put into effect targeted congestion mitigation strategies, and improve standard traffic float inside the state [6]. Furthermore, the findings of this observe will make contributions to the prevailing body of understanding on traffic control in growing towns, especially the ones facing similar challenges as Lagos State. The insights received from this research may be applied to decorate traffic management practices in other city regions in Nigeria [7].

To this end, the look at seeks to deal with the vital want for correct site traffic depth detection in Lagos State the use of a Bayesian estimation model. By leveraging comprehensive records resources and advanced statistical strategies, the proposed model pursuits to provide actual-time and correct traffic facts, empowering site traffic control government and customers to make knowledgeable choices and alleviate congestion efficaciously [8]. This study contributes to the field of site traffic management and urban planning and affords precious insights for addressing traffic demanding situations in Lagos State and similar city regions.

1.1 Statement of the problem

The lack of accurate and well-timed traffic records poses a considerable mission in dealing with site traffic congestion in Lagos State. Existing strategies rely on confined facts assets and simplistic models that fail to capture the complexities of traffic dynamics in a congested city environment. As a result, traffic control government warfare to make informed choices and put into effect proactive measures to relieve congestion [9]. There is a crucial want for a sophisticated estimation model that leverages complete information sources and includes advanced statistical techniques to provide accurate and real-time traffic depth statistics for powerful congestion management.

2. LITERATURE REVIEW

Traffic congestion is a vast trouble in city areas, impacting financial boom, productivity, and the environment. Various studies have targeted on estimating and quantifying the socioeconomical impact of congestion. The study [10] conducted a systematic literature review on the causes, effects, and techniques to minimize traffic congestion. The study [11] proposed a complex algorithm that considers factors which includes time cost, gasoline consumption, and CO₂ emissions to evaluate the socioeconomic impact of congestion. A comprehensive evaluation of site traffic prediction literature, highlighting studies procedures and developments in traffic glide prediction was conducted in the study [12] with the objective of identifying key trends and methodologies in the field. Consequently, traffic intensity detection is a key component in traffic monitoring, and intersection traffic state calculation. Different techniques are used for detecting site traffic depth, including using movement detection sensors in clever avenue lighting fixtures systems [13], laptop imaginative and prescient-based structures the use of traffic surveillance cameras [14], and traffic parameter mixtures in intersection detection gadgets [15]. These techniques' purpose to correctly decide the variety of cars or passengers present in a given place throughout a specific term.

In addition, previous researchers have proven the usage of system learning algorithms to acquire excessive accuracy in predicting specific degrees of traffic congestion. And beautify the accuracy and performance of traffic intensity detection systems. Traffic depth detection the use of system mastering has been a popular research area, where accurate and timely data approximately cutting-edge and forecast parameters of traffic flows can improve the performance of using delivery infrastructure [16]. Various system getting to know strategies and fashions have been proposed to come across and predict traffic depth, such as using convolutional neural networks (CNN) for powerful detection of site traffic congestion [17].

Additionally, deep studying fashions have been developed for traffic car detection, which could expect the varieties of motors the use of strategies together with body difference and contour mapping [18, 19]. These models aim to help in shooting and reading traffic data from CCTV surveillance in highways. Overall, system learning and deep getting to know strategies have shown promise in detecting and studying traffic intensity for numerous packages, consisting of traffic congestion manipulate and traffic surveillance.

2.1 Overview of Bayesian estimation

Bayesian estimation of traffic intensity is a statistical technique for estimating the ratio of the advent rate to the service fee in a queueing system. The Bayesian method uses prior statistics approximately the site traffic depth to update the posterior distribution of the traffic depth given the located information. This can be used to reap greater accurate estimates of the site traffic depth than conventional methods, consisting of maximum chance estimation. Bayesian estimation is a way of revising ideals considering observed statistics. It does not produce a single parameter estimate, however instead a distribution over the possible parameter values [9]. Bayesian estimation is based totally at the Bayesian interpretation of chance, in which probability expresses a degree of perception in an occasion [2]. The degree of belief may be primarily based on previous knowledge or personal ideals about the event.

To perform Bayesian estimation, we want to specify a prior distribution that displays our initial beliefs approximately the parameter of hobby. Then, we use Bayes' theorem to update our ideals based on the information we observe. The posterior distribution, or the conditional distribution of the parameter given the records, can be computed the usage of the Bayes theorem [20]. The facts from the statistics and our preceding ideals are combined to create the posterior distribution.

Bayesian estimation can be applied to various types of data and models. For example, we can use Bayesian estimation to estimate the proportion of right-handed people in a population, or the mean height of a certain species of plants, or the effect of a treatment on a disease outcome. In each case, we need to choose an appropriate prior distribution and a likelihood function that models how the data are generated given the parameter value [6].

Bayesian estimation has some advantages over other methods of inference, such as frequents statistics where by Bayesian approach offers advantages in traffic intensity detection by allowing for the incorporation of prior knowledge, handling uncertainty, integrating multiple data sources, adapting to changing conditions, and providing robust estimates in the presence of outliers or limited data. These benefits make Bayesian methods to be well-suited for applications where accurate and reliable estimation of traffic intensity is highly critical.

For example, Bayesian estimation allows us to incorporate prior knowledge and uncertainty into our analysis, this is particularly useful because in real-world scenarios, we often have some idea about the typical patterns of traffic intensity based on historical data or domain expertise. By integrating this prior know-how into the Bayesian version, we are able to improve the accuracy of traffic depth detection, it additionally inference can offer dependable estimates even if information is limited or noisy. This is beneficial in traffic intensity detection in which records may not continually be abundant or

may be suffering from various factors such as climate, events, or road conditions. The Bayesian framework permits for updating ideals as new information becomes to be had, that can help in converting visitors' conditions. Bayesian models also can adapt and research over time as extra information is amassed. This adaptability is vital in traffic depth detection where styles can also trade through the years [16]. By continuously updating the version with new statistics, Bayesian methods can offer greater accurate and updated estimates of traffic depth. Even observed you however you and to make probabilistic statements about the parameter without counting on repeated sampling or self-belief durations. Bayesian estimation additionally offers an herbal manner to evaluate different fashions or hypotheses using the Bayes issue, that's the ratio of posterior chances for 2 models [4].

However, Bayesian estimation also has a few challenges and barriers. For instance, selecting a previous distribution may be subjective and controversial, and might have an effect on the outcomes of the analysis. Also, computing the posterior distribution may be computationally extensive and require specialized algorithms, which includes Markov chain Monte Carlo (MCMC) strategies [2]. In summary, taking pictures spatial and temporal variations in Bayesian models for traffic depth detection entails structuring the version to explicitly account for geographical modifications over time. By utilizing suitable priors, hierarchical systems, and dynamic modeling techniques, Bayesian strategies can efficaciously and correctly capture the complicated dynamics of site visitors' depth across each space and time in Lagos State.

Furthermore, decoding and speaking Bayesian consequences may be tough for some audiences who aren't acquainted with the Bayesian framework.

2.2 Bayesian estimation model

A direct and powerful class approach based totally at the Bayes theorem is referred to as Naive Bayes. It makes the idea that, given the magnificence name, the facts' capabilities are conditionally impartial. Stated in another way, it presupposes that a function's effect at the magnificence is ignorant of other functions. In data technological know-how and gadget learning, the Naive Bayes set of rules is a famous, honest technique. It's an algorithm for classifiers that makes use of possibility.

It also the inspiration of this version is Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

where, A and B are any events, $P(A|B)$ is the likelihood of A conditionally given B.

To implement Naive Bayes in Python, we can use the scikit-learn library [21], which provides various implementations of Naive Bayes for different types of features. In this project, we consider the Gaussian Naive Bayes for numerical features with Gaussian distribution, and multinomial Naive Bayes for discrete features with multinomial distribution.

2.2.1 Gaussian Naive Bayes' classification model

In the context of classification, we can use Bayes' theorem to calculate the probability of a class label C_k .

For each class C_k , estimate μ_{ik} and variance of σ_{ik}^2 of each feature x_i , assuming it follows a Gaussian distribution given a

feature vector $x = (x_1, x_2, \dots, x_n)$ as follows:

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)} \quad (2)$$

where, $P(x|C_k)$ is the probability of the characteristics given the class, $P(C_k)$ is the class's prior probability, and $P(x)$ provides proof of the characteristics' marginal probability.

To make a prediction for a new instance x , we use the following rule:

$$\hat{y} = \arg \max_k P(C_k|x) \quad (3)$$

Step 1: Calculate the prior probability (C_k) for each class (C_k)

Step 2: Use Bayes' theorem to calculate the posterior probability $P(x|C_k)$

That is, we choose the class that has the highest probability given the features. However, calculating $P(x|C_k)$ and $P(x)$ can be difficult, particularly if there are a lot of features. The Gaussian assumption is useful in this situation [22]. We can simplify the likelihood as follows by assuming that the features are conditionally independent given the class:

$$P(x|C_k) = P(x_1|C_k)P(x_2|C_k) \dots P(x_n|C_k) = \prod_{i=1}^n P(x_i|C_k) \quad (4)$$

where, $P(C_k)$ is the prior probability of C_k , $P(x|C_k)$ is the likelihood of feature x_i given class C_k for each $P(x_i|C_k)$ is a Gaussian probability density function with μ_{ik} and standard deviation σ_{ik} :

$$P(x_i|C_k) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} e^{-\frac{(x_i-\mu_{ik})^2}{2\sigma_{ik}^2}} \quad (5)$$

This means that we only need to estimate the probability of each feature given the class, which can be done using various methods depending on the type of the feature (e.g., categorical, or numerical). Similarly, we can ignore $P(x)$ in the denominator since it is constant for all classes. As a result, the posterior probability may be rewritten as:

$$P(C_k|x) \propto P(C_k) \prod_{i=1}^n P(x_i|C_k) \quad (6)$$

Step 3: Assign the class label \hat{y} for a new instance x by setting the class C_k with the highest posterior probability $P(x|C_k)$

2.2.2 Multinomial Naive Bayes classification model

When it comes to classification, the Bayes theorem allows us to determine the likelihood of a class label C_k .

For each class C_k , estimate the probability $P(C_k|x)$ of each term x_i occurring, given a feature vector $x = (x_1, x_2, \dots, x_n)$ as follows:

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)} \quad (7)$$

where, $P(x|C_k)$ is the likelihood of the features given the class, $P(C_k)$ is the prior probability of the class, and $P(x)$ is the

evidence of marginal probability of the features [23].

To make a prediction for a new instance x , we use the following rule:

$$\hat{y} = \arg \max_k P(C_k|x) \quad (8)$$

Count (x_i in C_k) is the number of terms x_i appears in documents of class C_k .

α is a smoothing parameter (often set to 1, Laplace smoothing to handle terms not present in training data.

Step 1: calculate the prior probability $P(C_k)$ for each class C_k

Step 2: use Bayes' theorem to calculate the posterior probability $P(C_k|x)$

That is, we choose the class that has the highest probability given the features. However, calculating $P(x|C_k)$ and $P(x)$ can be challenging, especially when the number of features is large. This is where the multinomial assumption comes in handy. By assuming that the features are conditionally independent given the class, we can simplify the likelihood as follows:

$$\begin{aligned} P(x|C_k) &= P(x_1|C_k)P(x_2|C_k) \dots P(x_n|C_k) \\ &= \prod_{i=1}^n P(x_i|C_k) \end{aligned} \quad (9)$$

where, each $P(x_i|C_k)$ is a multinomial probability function with parameter θ_{ik} :

$$P(x_i|C_k) = \binom{n}{x_i} \theta_{ik}^{x_i} (1 - \theta_{ik})^{n-x_i} \quad (10)$$

where, n is the total number of trials (e.g., words or pixels), and x_i is the number of successes (e.g., occurrences or intensities) for feature i . The parameter θ_{ik} is the probability of success for feature i given class k , and it can be estimated by:

$$\hat{\theta}_{ik} = \frac{N_{ik} + \alpha}{N_k + n\alpha} \quad (11)$$

where, N_{ik} is the number of times feature i appears in class k , N_k is the total count of all features in class k , and α is a smoothing parameter that prevents zero probabilities.

Similarly, we can ignore $P(x)$ in the denominator since it is constant for all classes. Therefore, we can rewrite the posterior probability as:

$$P(C_k|x) \propto P(C_k) \prod_{i=1}^n P(x_i|C_k) \quad (12)$$

Step 3: assign the class label \hat{y} for a new document x by selecting the class C_k with the highest posterior probability $P(x_i|C_k)$.

2.2.3 Simple statement of Bayes' theorem

In its most basic setting, which involves only discrete distributions, the Bayes theorem links the conditional and marginal probabilities of event A and B as a non-vanishing probability. A specific scenario involving discrete probability distribution of data and continuous prior and posterior probability distributions was presented by Thomas Bayes [12]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (13)$$

Every term in the Bayes theorem has a common name:

- (1) $P(A)$ is either A's marginal probability or prior probability. It is "prior" in the sense that it disregards any knowledge that B may have.
- (2) $P(A|B)$ is the likelihood of A given B, contingent on. Since it is based on or derives from the given value of B, it is also known as the posterior probability.
- (3) $P(B|A)$ is the probability of B given A, conditionally. Another name for it is the chance.
- (4) $P(B)$ is used as a normalization constant and represents the marginal or prior probability of B.

The mathematical illustration provided by the Bayes theorem illustrates how the conditional probability of B given A [24].

2.2.4 Bayes' extension for two or more events

Bayes-like theorems encompass more than two events. As an instance:

$$P(A|B \cap C) = \frac{P(A)P(B|A)P(C|A \cap B)}{P(B)P(C|B)} \quad (14)$$

A few simple steps will get you there from the definition of conditional probability and the Bayes theorem:

$$\begin{aligned} P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= \frac{P(C|A \cap B)P(A \cap B)}{P(B)P(C|B)} = \frac{P(A)P(B|A)P(C|A \cap B)}{P(B)P(C|B)} \end{aligned} \quad (15)$$

Similarly,

$$P(A|B \cap C) = \frac{P(B|A \cap C)P(A|C)}{P(B|C)} \quad (16)$$

can be derived in the following way and is thought of as a conditional Bayes theorem:

$$\begin{aligned} P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= \frac{P(B|A \cap C)P(A|C)P(C)}{P(C)P(B|C)} = \frac{P(B|A \cap C)P(A|C)}{P(B|C)} \end{aligned} \quad (17)$$

A common method is to work with a decomposition of the joint possibility and marginalizing (integrating or summarizing) over the variables that are not of interest. Making use of this function would possibly drastically reduce the calculations. Depending at the form of decomposition, it can be smooth to show that a few integrals need to be 1 so as for the decomposition to terminate [25]. For example, a Bayesian community specifies the factorization of a joint distribution of several variables, wherein the conditional possibility of every variable given the opposite variables takes a fairly easy form.

3. RELATED WORKS

This study utilized Generative Adversarial Network (GAN) Sample Enhancement to decide site visitors' depth. First off,

statistics consolidation, lacking price processing, and records normalization are the primary duties of the facts preprocessing module. Malicious pattern introduction based totally on GAN module is dubbed in this foundation. Data is constantly generated and discriminated for the input noise using GAN. In order to optimize the era model and discriminant model parameters, the discriminant loss-which might also quantify the distinction among generative malicious samples and unique malicious samples-is backpropagated [26]. When the discriminant loss reaches the predetermined level, backpropagation is terminated. Now that the generative malicious samples match the original malicious samples, the imbalance inside the PIoT traffic facts is resolved and there are extra minority samples. The difficulty of an over-reliance on professional experience in function selection is then resolved with the aid of serially applying variance evaluation and correlation analysis to create a malicious site visitors function set, considering the divergence of capabilities and the Pearson correlation coefficient among capabilities and the category. Finally, using the Naive Bayes set of rules module, the Naive Bayes classifier is skilled to understand distinctive forms of dangerous traffic in malicious traffic detection as shown in Figure 1 [6].

The study [27] presents a monitoring system developed at Politecnico di Torino, initially designed for freight vehicles to detect accelerations and temperatures for diagnostics. Enhanced to include vehicle braking system diagnostics, the system identifies braking anomalies to aid maintenance. It outlines essential features for modern monitoring systems on railway freight wagons and introduces a new version tested on

a scaled roller-rig to simulate abnormal braking conditions safely. Equipped with an autonomous axle generator, diagnostic processing capabilities, and communication protocols, the system ensures reliable operation and efficient transmission of diagnostic information.

3.1 Evaluations of the rate of delays at different intersections (traffic count analysis)

The study focused on the traffic congestion problem along the Ipaja-Ikotun road corridor in Lagos Metropolis, Nigeria. The paper begins with an introduction that highlights the importance of addressing traffic congestion issues in urban areas. The problem identified the high rate of delay and monetary cost of delays caused by traffic congestion in the study area. The study's objectives are to compile comprehensive data on the current state of traffic, pinpoint issues, and demonstrate the financial impact of delays on the economy [28].

The study's findings in Table 1 demonstrate that the road corridor's traffic flow issue is a result of lost man-hours in traffic. The report urges the government to immediately undertake road renovation, particularly for the road arteries that provide detours for the majority of the study area's traffic. It is also advised that the Iyana-Ipaja Motor Park Garage be moved in order to allow unhindered traffic movement along the road corridor and reduce the number of man-hours lost to traffic [28]. The study comes to the realization that traffic congestion problems get worse as the populace grows and greater humans turn out to be wealthy and use era.

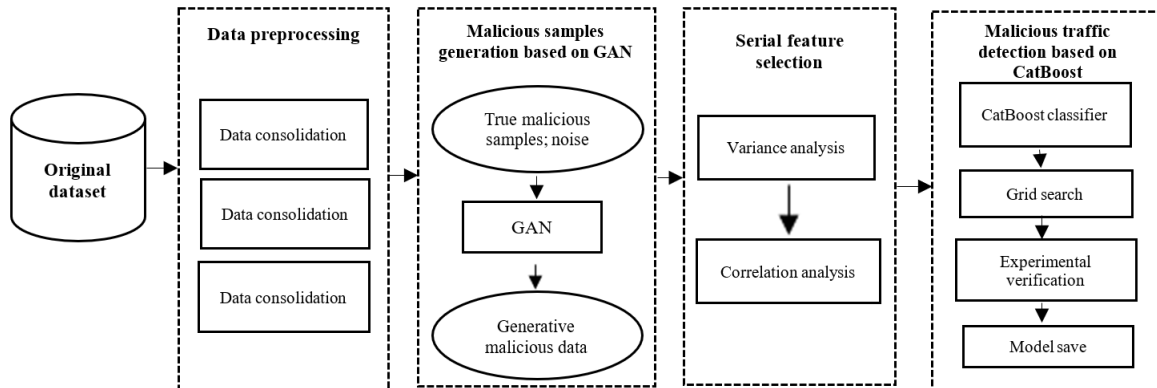


Figure 1. A diagram showing GAN sample enhancement [26]

Table 1. Average daily vehicle characteristics in zone A and zone B [28]

Vehicle Type	A1 Day 1	A2 Day 2	A3 Day 3	A0 Average	B1 Day 1	B2 Day 2	B3 Day 3	B0 Average
Cars	6773	6416	6930	6706	2451	5515	6046	4671
Motorcycle	551	243	418	404	341	417	2386	1048
Tricycle	7924	914	4213	4350	1487	1986	1935	1803
Buses	5202	3020	5746	4656	2165	5462	3451	3692
HDV's	55	36	42	44	21	41	61	41
BRT	41	32	38	37	34	42	49	42
Bicycles	13	9	10	11	16	14	18	16
Luxury	36	22	31	30	19	23	32	25
Mini Bus	2668	1957	3758	2794	1256	3925	3561	2914
Total	23263	12649	21186	19032	7790	17425	17539	14251

3.2 Traffic intensity in M/M/1 queue under precautionary loss function

This study centred on the utility of Bayesian estimators

within the context of traffic intensity estimation in M/M/1 queuing structures. The creation affords an overview of the significance of accurate traffic depth estimation and the constraints of traditional estimation strategies. The problem

highlighted the need for a more robust estimation approach that considers both overestimation and underestimation [5].

The paper aims to offer a new weighted square error loss function for Bayesian estimators that allocates equal losses to under- and overestimations as shown in Table 2.

The results as shown in Table 3, contains how well the cautioned weighted rectangular error loss function works to increase site visitors' depth estimation accuracy. The suggested technique is prime when compared to Bayesian estimators and maximum chance (ML) below a precautionary loss feature.

In conclusion, the study highlights the importance of Bayesian estimators in site visitors' intensity estimation and the benefits of the usage of a weighted rectangular blunders loss feature [4].

Table 2. Estimators and squared error of estimators with $\rho=0.5, n=50$ [5]

C	P ₁	ER(p ₁)	A	B	P ₂	ER(p ₂)
0.0	0.4967	0.0023	1.0	1.0	0.4967	0.0023
1.0	0.4917	0.0025	1.0	1.5	0.4943	0.0023
1.5	0.4891	0.0026	1.0	2.0	0.4919	0.0024
2.0	0.4865	0.0027	1.5	1.0	0.4992	0.0022
2.5	0.4839	0.0028	2.0	1.5	0.4993	0.0022
3.0	0.4812	0.0029	2.0	3.0	0.4922	0.0023

Table 3. Estimators and squared error of estimators with $\rho=0.8, n=50$ [4]

C	P ₁	ER(p ₁)	A	B	P ₂	ER(p ₂)
0	0.7956	6.8898e004	1.0	1.0	0.7956	6.8898e004
1.0	0.7948	7.0810e004	1.0	1.5	0.7940	7.1308e004
1.5	0.7943	7.1830e004	1.0	2.0	0.7924	7.4208e004
2.0	0.7939	7.2894e004	1.5	1.0	0.7960	6.8004e004
2.5	0.7935	7.30803004	2.0	1.5	0.7948	6.9286e004
3.0	0.7931	7.5153e004	2.0	3.0	0.7901	7.8587e004

4. METHODOLOGY

The methodology of this research consists of four main steps: data collection, data preprocessing, data analysis, and data evaluation. Figure 2 shows the model architecture for data from the input data, the processing stage, to the output.

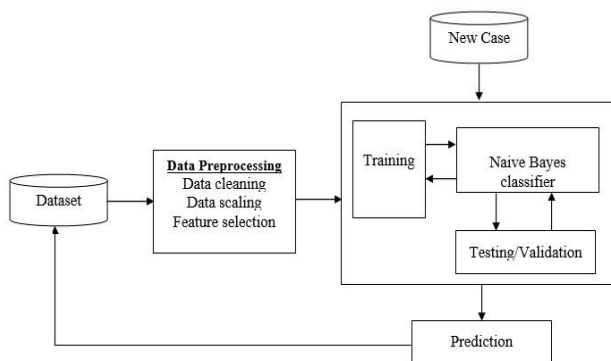


Figure 2. Model architecture

The following is a brief overview of each step:

Data collection: This step involves obtaining the secondary data from Kaggle repository, which contains 48120 observations of the number of vehicles intersecting at four

junctions in Lagos State at various times of the day and on different days of the week [7]. The data was collected by observing the number of vehicles passing through each junction using cameras or sensors. The data was obtained from <https://www.kaggle.com/ahmedlahlou/accidents-in-france-from-2015-to-2017>.

Data preprocessing: This step involves ensuring the quality and suitability of the data for further analysis by handling missing values, outliers, and inconsistencies. An Exploratory Data Analysis was also carried out to identify any data quality issues, missing values, outliers, or patterns within the dataset [20]. Additionally, data transformation was used, which included turning unstructured data into characteristics that made sense and enhanced the functionality of machine learning models. The following were extracted: the hours, the day of the week, month, and year, generating cyclic features for time of day, or calculating time differences between events. Data visualization provided intuitive insights into traffic patterns, data quality, and aid in decision-making for subsequent analysis steps. All these were subject to Feature selection, Data scaling, cross-validation and data splitting [8].

Data analysis: This step involves applying the Bayesian Estimation Model to estimate the traffic intensity in Lagos State. The Gaussian Naive Bayes classifier and the multinomial Naive Bayes classifier are the models chosen for this project. These are certain kinds of Bayesian Estimation Models that use the Bayes theorem under the mistaken presumption that each pair of features provided the class variable will always be conditionally independent [29]. The features are assumed to be regularly distributed by the Gaussian Naive Bayes classifier, whereas they are assumed to be multinomially distributed by the multinomial Naive Bayes classifier. These models were chosen because they are simple, scalable, and suitable for high-dimensional data [25].

A Gaussian distribution is assumed to easily estimate the probabilities for input variables using the Gaussian Probability Density Function. The GaussianNB function of the sklearn Naive_Bayes class in the scikit-learn library of Python was used to implement the model. The data will be divided into 80% for the training and the remaining 20% for testing and validation.

Data evaluation: This step entails evaluating the performance and accuracy of the Bayesian Estimation Model for visitors' intensity detection in Lagos State. An Evaluation Metrics was completed in understanding the version's accuracy, precision, recall, and standard overall performance, allowing effective model assessment and contrast [30]. Accuracy gauges how frequently the version makes accurate predictions, precision determines the percentage of wonderful predictions that the model as it should be predicting, recall evaluates the frequency with which the model properly predicts among fine cases, and F1-rating calculates the precision and consider harmonic suggest. These metrics have been calculated the use of confusion matrices and type reviews. The fashions were additionally compared the usage of the Confusion Matrices, which measure how well the version can distinguish between exclusive lessons [20].

4.1 Evaluations metrics

These metrics collectively help in knowledge the version's accuracy, precision, recollect, and ordinary overall performance, allowing powerful model assessment and contrast [31].

(1) Accuracy score: The accuracy rating is a typically used assessment metric for category fashions. It calculates the proportion of the model's total wide variety of predictions that are correct. It gives a large image of the version's effectiveness, but while running with unbalanced datasets, it couldn't be sincere.

(2) Classification report: The class record gives an intensive assessment of the effectiveness of a classification model. Metrics like remember, accuracy, F1 rating, and assistance are included for every class. The F1 rating is the harmonic mean of precision and remember. Precision gauges the accuracy of fantastic predictions, consider gauges the capability to understand nice cases. The quantity of occurrences of every magnificence within the dataset is called the aid.

(3) Confusion matrix: A confusion matrix is a table that visualizes the performance of a category version by way of summarizing the results of predictions. Its presentations the quantity of true positives, proper negatives, fake positives, and false negatives by evaluating the projected training with the real instructions. It enables to perceive the version's overall performance in phrases of efficiently and incorrectly labelled times for each magnificence. The Gaussian Naive Bayes classifier's anticipated elegance labels (y_{pred}) and real class labels (y_{test}) had been used to calculate the confusion matrix. The confusion matrix became then visualized the use of a confusion matrix Display item, imparting a graphical representation of the distribution of predicted and real elegance labels. The plot displayed the confusion matrix, and the name "Confusion Matrix Distribution of Gaussian Naive Bayes" became assigned to the plot.

5. RESULTS AND ANALYSIS

5.1 Training, testing accuracy score, precision and recall score for Gaussian Naive Bayes

From Figure 3, It indicates that the Gaussian Naive Bayes classifier's education accuracy rating became 0.96, that means that the model became rather correct in its prediction of the education set. Similarly, the testing accuracy rating was additionally 0.96, suggesting that the version achieved properly on unseen records as well. The precision and recollect ratings for Gaussian Naive Bayes have been both 0.96, as displayed inside the bar plot. These ratings indicated that the version had an excessive precision and bear in mind price, reflecting its ability to accurately identify and classify visitors' times efficaciously.

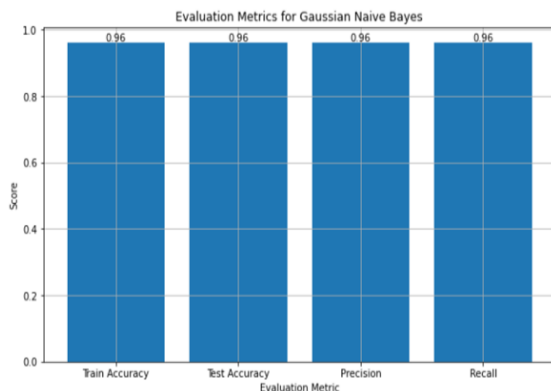


Figure 3. Bar chart of the evaluation metric for Gaussian Naive Bayes

Overall, the Gaussian Naive Bayes classifier tested robust performance in phrases of accuracy, precision, and consider, making it an appropriate version for visitor's detection tasks.

5.2 Classification report for Gaussian Naive Bayes

In Table 4, the classification record gives a radical assessment of the Gaussian Naive Bayes classifier's performance within the visitor's detection job.

Table 4. Classification report metric for Gaussian Naive Bayes

Class	Precision	Recall	F1-Score	Support
0	0.96	0.97	0.97	3162
1	0.92	0.95	0.94	2948
2	0.99	0.96	0.97	3514
Accuracy			0.96	9624
Macro avg	0.96	0.96	0.96	9624
Weighted avg	0.96	0.96	0.96	9624

Gaussian Naive Bayes accuracy: 95.97%

For class 0 (negative), with a precision of 0.96, the classifier successfully recognized 96% of the instances that were predicted to be negative. The recall rating became 0.97, indicating the proportion of real poor cases that had been well detected. The recall score, which measures the proportion of actual negative instances correctly identified, was 0.97. The F1-score, a harmonic mean of precision and recall, was 0.97. For magnificence 1 (impartial), the precision score changed into 0.92, indicating that 92% of the times predicted as neutral were successfully classified. The remember score for this magnificence turned into 0.95. The F1-rating turned into 0.94. For class 2 (nice), the precision rating become 0.99, indicating that 99% of the times expected as high-quality were correctly categorized. The keep in mind rating for this class become 0.96, and the F1-score changed into 0.97.

The Gaussian Naive Bayes classifier's total accuracy became 0.96, indicating that it classified the site traffic instances with an excessive degree of accuracy. Additionally, the F1-score, don't forget, and precision macro averages have been all 0.96, suggesting that each one classes completed in addition. Taking into consideration the elegance imbalance, the weighted common came in at 0.96. In end, the Gaussian Naive Bayes classifier validated wonderful overall performance throughout all classes, with excessive precision, remember, and F1-score. It is a reliable model for traffic detection duties.

5.3 Confusion matrix for Gaussian Naive Bayes

In Figure 4 and Table 5, the Gaussian Naive Bayes classifier's overall performance at the visitor's identity project is visually represented with the aid of the confusion matrix. It displays the quantity of times labelled into each elegance and compares them to the real class labels.

For class 0 (negative), the classifier efficaciously categorised 3072 instances as bad and misclassified 90 times as neutral. However, it misclassified 116 instances as negative and 37 instances as positive.

For class 2 (positive), the classifier correctly classified 3369 instances as positive. It misclassified 145 instances as neutral, but there were no instances misclassified as negative.

Overall, the Gaussian Naive Bayes classifier demonstrated

strong performance in correctly classifying the majority of instances across all classes. However, it had a higher tendency to misclassify instances from class 1 (neutral) as negative or positive. These misclassifications can be further analysed and improved to enhance the model's accuracy.

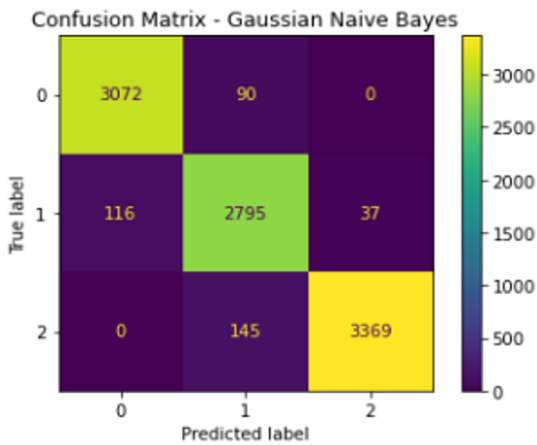


Figure 4. Confusion matrix for Gaussian Naive Bayes

Table 5. Tabular representation of confusion matrix for Gaussian Naive Bayes

	Negative	Neutral	Positive
Negative	3072	90	0
Neutral	116	2795	37
Positive	0	145	3369

5.4 Multinomial Naive Bayes results

Based on the visual representation and evaluation metrics in Figure 5, the multinomial Naive Bayes classifier achieved a training accuracy score of 0.69 and a test accuracy score of 0.69. These scores indicated the proportion of correctly classified instances in both the training and test datasets. The classifier's precision score is 0.67, indicating that the model can correctly identify positive cases. Recall, which measures how well the model captures positive instances from the dataset, was 0.69.

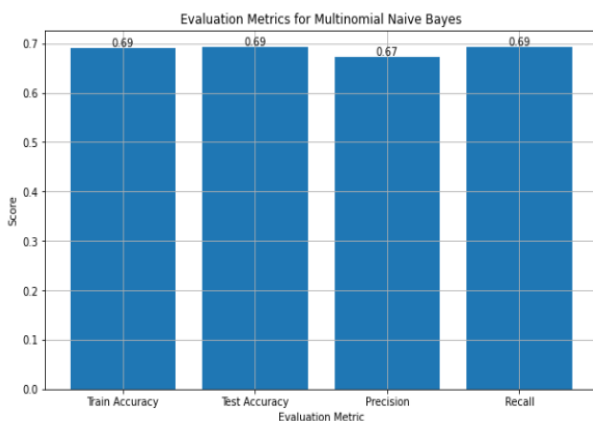


Figure 5. Bar chart of the evaluation metric for multinomial Naive Bayes

The bar plot visualization provides a clear overview of these performance metrics, highlighting the accuracy, precision, and recall scores for the multinomial Naive Bayes classifier. These results indicated that the classifier performs reasonably well,

but there was room for improvement, particularly in precision, to enhance the accuracy of positive class predictions.

5.5 Multinomial Naive Bayes classification report

In Table 6, the classification report supplied an in-depth evaluation of the multinomial Naive Bayes version's performance. The accuracy score for the version is 0.69, indicating the general share of successfully categorized instances. Based at the precision rankings, the model finished the subsequent duties with 0.70 precision for class 0, 0.55 precision for sophistication 1, and 0.75 precision for class 2. The precision of a version is its capability to efficiently categorize instances for every magnificence.

Table 6. Classification report for multinomial Naive Bayes

Class	Precision	Recall	F1-Score	Support
0	0.70	0.83	0.76	3162
1	0.55	0.34	0.42	2948
2	0.75	0.86	0.80	3514
Accuracy			0.69	9624
Macro avg	0.67	0.68	0.66	9624
Weighted avg	0.67	0.69	0.67	9624
Multinomial Naive Bayes accuracy: 69.22%				

The version produced bear in mind values of 0.83 for class 0, 0.34 for sophistication 1, and 0.86 for sophistication 2. Recall reflected the model's capacity to successfully discover times of every elegance from the dataset.

Class 0, class 1, and class 2 have f1-scores of 0.76, 0.42, and 0.80, respectively, which integrate precision with remember. Indicating the version's average overall performance throughout all training, the weighted average f1-score turned into 0.67.

A thorough evaluation of the multinomial Naive Bayes model's effectiveness in categorizing the traffic information become made feasible via the classification file, which offered insightful records at the version's precision, take into account, and f1-rating for every magnificence.

5.6 Confusion matrix-multinomial Naive Bayes

In Figure 6 and Table 7, the confusion matrix furnished an overview of the multinomial Naive Bayes version's performance in classifying the traffic records. It is composed of expected class labels inside the columns and real magnificence labels in the rows.

For the "Negative" class, the model correctly predicted 2639 times, misclassified 445 instances as "Neutral", and misclassified 78 instances as "Positive".

In the "Neutral" magnificence, the version successfully predicted 993 instances, misclassified 1029 times as "Negative", and misclassified 926 times as "Positive".

In the "Positive" magnificence, the model correctly predicted 3030 times, misclassified a hundred and five instances as "Negative", and misclassified 379 instances as "Neutral".

A visible depiction of the model's performance in terms of effectively categorised cases and misclassifications throughout numerous classes turned into made to be had via the confusion matrix. It helped to pick out the areas where the version might also struggle and the classes that had been at risk of misclassification.

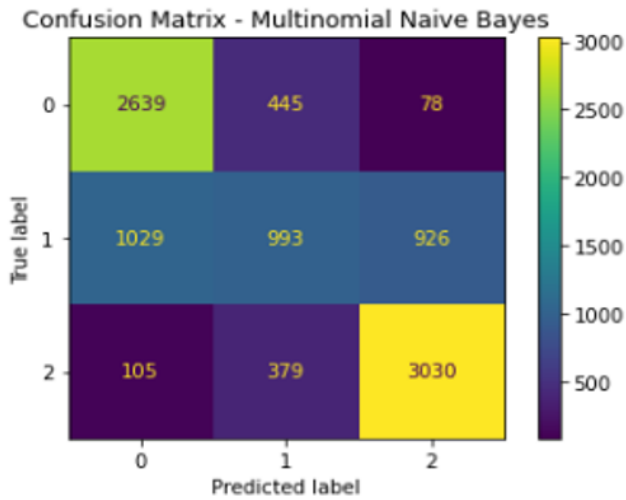


Figure 6. Confusion matrix for multinomial Naive Bayes

Table 7. Tabular representation of confusion matrix for multinomial Naive Bayes

	Negative	Neutral	Positive
Negative	2639	445	78
Neutral	1029	993	926
Positive	105	379	3030

5.7 Performance comparison: Gaussian Naive Bayes vs. multinomial Naive Bayes

In comparing the overall performance of the Gaussian Naive Bayes and multinomial Naive Bayes models on the traffic detection mission, several key metrics have been considered: educate accuracy, check accuracy, precision, recall, and the confusion matrix as shown in Table 8.

Table 8. Tabular representation performance comparison of the two models

Metric	Gaussian Naive Bayes	Multinomial Naive Bayes
Train Accuracy	0.959658	0.68989
Test Accuracy	0.959684	0.692228
Precession	0.960321	0.671463
Recall	0.959684	0.692228

The underneath is represented in Table 8 except the confusion matrix:

Train Accuracy: The Gaussian Naive Bayes model carried out a train accuracy of 95.97%, indicating that it accurately classified 95.97% of the education instances. The multinomial Naive Bayes model finished a train accuracy of 68.99%.

Test Accuracy: Both models exhibited comparable performance in terms of test accuracy. The multinomial Naive Bayes model obtained a touch better test accuracy of 69.22%, compared to the 95.97% finished via the Gaussian Naive Bayes model.

Precision: Precision refers back to the capacity of the model to efficaciously identify instances of a selected class. The Gaussian Naive Bayes model carried out a precision score of 96.03%, indicating its ability to correctly classify instances. The multinomial Naive Bayes version accomplished a precision score of 67.15%.

Recall: Recall measures the model's ability to correctly

identify instances of a specific class out of all the instances belonging to that class. The Gaussian Naive Bayes model achieved a recall score of 95.97%, indicating its effectiveness in identifying relevant instances. The multinomial Naive Bayes model achieved a recall score of 69.22%.

Confusion Matrix: The confusion matrix affords an in-depth breakdown of the model's predictions for each magnificence. For the Gaussian Naive Bayes version, the confusion matrix showed that it effectively categorized 3,072 times as negative, 2,795 instances as neutral, and three,369 instances as advantageous. Similarly, the multinomial Naive Bayes version successfully categorised 2,639 instances as bad, 993 instances as neutral, and 3,030 instances as high-quality.

6. CHALLENGES FACED IN LAGOS STATE, NIGERIA

As Nigeria's economic center, Lagos attracts humans from all over the country in search of employment possibilities, further growing the population and the variety of automobiles on the road. Lagos State, faces sizeable challenges with visitors' congestion because of several factors, along with fast populace growth, infrastructure boundaries, and socioeconomic situations. Lagos, with a populace estimated to be over 20 million human beings and awareness of industrial activities in areas like Victoria Island, Ikeja, and Lekki. This fast urbanization places titanic stress on existing transportation infrastructure specially at some point of peak hours as human beings go back and forth to and from paintings. The street community in Lagos is insufficient to deal with the extent of visitors as roads are narrow, poorly maintained, lack proper drainage structures, and restrained alternative routes leading to common site visitors' jams. The city is predicated heavily on casual transport structures like minibuses (danfos) and bikes (okadas), which can be chaotic and make contributions to congestion. Hence, because of insufficient public shipping, many citizens rely on non-public motors, increasing the number of motors on the street.

Many traffic alerts used for traffic control are old or not synchronized, leading to inefficient traffic glide. Traffic laws aren't continuously enforced, main to troubles like unlawful parking, avenue buying and selling, and disregard for site visitors' guidelines. Lack of Real-Time Traffic Data hinders powerful site visitors' management and well-timed response to incidents. Delays and inefficiencies in implementing transportation policies and projects further exacerbate congestion problems. Addressing these challenges requires a comprehensive approach, including significant investments in infrastructure, improved public transportation systems, efficient traffic management, and better urban planning practices.

7. PRACTICAL IMPLICATIONS OF THE RESEARCH FINDINGS

Accurate real-time traffic information provided by the model can be used by traffic management authorities to mitigate congestion by allowing traffic signals to adapt to current traffic conditions, optimizing green light durations to alleviate congestion, synchronizing traffic lights along main corridors to create green waves, reducing stop-and-go driving and improving traffic flow. Providing drivers with up-to-date

traffic information through mobile apps, websites, and roadside displays which helps them make informed decisions about their routes. Accurate real-time traffic information also provides urban planners to modify public transit routes and schedules based on traffic patterns to improve efficiency and attractiveness of public transport, use traffic data to identify bottlenecks and high-traffic areas, guiding the development of new roads, bridges, or transit systems lastly use traffic data to assess and mitigate the environmental impact of congestion, such as emissions and noise pollution.

Hence, by utilizing accurate real-time traffic information provided by the model, both traffic management authorities and urban planners can implement targeted and efficient strategies to reduce congestion, improve traffic flow, and enhance the overall transportation experience.

8. CONCLUSION AND RECOMMENDATION

In conclusion, this research work successfully applied Bayesian Estimation Models to detect traffic intensity in Lagos State. Through the implementation of Naive Bayes and other Bayesian Estimation techniques, a significant level of accuracy was achieved in classifying traffic intensity levels. The models demonstrated their capability to capture patterns and trends, providing valuable insights for traffic management and planning.

The findings of this research indicate that Bayesian Estimation Models have great potential in enhancing traffic management systems in Lagos State. By utilizing historical traffic data and incorporating real-time information, these models can assist in predicting traffic intensity, optimizing traffic flow, and making informed decisions for efficient transportation planning.

REFERENCES

- [1] Schölkopf, B., Platt, J., Hofmann, T. (2007). Advances in neural information processing systems. In Proceedings of the 2006 Neural Information Processing System Conference, Cambridge, Ma. <http://mitpress.mit.edu/9780262195683/>.
- [2] Nwaigwe, D.N., Amiara Chidiebere, A., Okwunze, C.F., Egege, C.C. (2019). Analytical study of causes, effects and remedies of traffic congestion in Nigeria: Case study of Lagos state. *International Journal of Engineering Research and Advanced Technology*, 5: 11-19. <https://doi.org/10.31695/ijerat.2019.3542>
- [3] Afrin, T., Yodo, N. (2020). A survey of road traffic congestion measures towards a sustainable and resilient transportation system. *Sustainability*, 12(11): 4660. <https://doi.org/10.3390/su12114660>
- [4] Arjas, E., Heikkinen, J. (1997). An algorithm for nonparametric Bayesian estimation of a Poisson intensity, <https://jukuri.luke.fi/handle/10024/506706>.
- [5] Ren, H.P., Li, J.P. (2012). Bayes estimation of traffic intensity in M/M/1 queue under a new weighted square error loss function. *Advanced Materials Research*, 485: 490-493. <https://doi.org/10.4028/www.scientific.net/AMR.485.490>
- [6] Ross, K. (2023). Chapter 5 introduction to estimation. In *An Introduction to Bayesian Reasoning and Methods*. https://bookdown.org/kevin_davisross/bayesian-reasoning-and-methods/estimation.html.
- [7] Almeida, M.A., Cruz, F.R. (2018). A note on Bayesian estimation of traffic intensity in single-server Markovian queues. *Communications in Statistics-Simulation and Computation*, 47(9): 2577-2586. <https://doi.org/10.1080/03610918.2017.1353614>
- [8] Zhao, Y. (2019). *Gaussian Processes for Positioning Using Radio Signal Strength Measurements*. Linköping University Electronic Press, Vol. 1968. <https://doi.org/10.3384/DISS.DIVA-153944>
- [9] Speekenbrink, M. (2023). Chapter 15 introduction to Bayesian estimation. In *Statistics: Data Analysis and Modelling*. <https://mspeekenbrink.github.io/sdam-book/ch-Bayes-estimation.html>.
- [10] Siegrist, K. (2020). 7.4: Bayesian estimation. https://stats.libretexts.org/Bookshelves/Probability_Theory/Probability_Mathematical_Statistics_and_Stochastic_Processes_%28Siegrist%29/07%3A_Point_Estimation/7.04%3A_Bayesian_Estimation.
- [11] Sharma, P. (2021). Implementing gaussian naive bayes in Python Sklearn. *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2021/11/implementation-of-gaussian-naive-bayes-in-python-sklearn/>.
- [12] GaussianNB, scikit-learn 0.22.1 documentation. https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html.
- [13] Sharma, P. (2021). Gaussian naive bayes implementation in Python Sklearn. <https://machinelearningknowledge.ai/gaussian-naive-bayes-implementation-in-python-sklearn/>.
- [14] Aquino-Santos, R., Edwards, A., Rangel-Licea, V. (2012). Wireless technologies in vehicular ad hoc networks: Present and future challenges. *IGI Global*. <https://doi.org/10.4018/978-1-4666-0209-0>
- [15] Etz, A. (2016). Sunday Bayes: A brief history of Bayesian stats. <https://alexanderetz.com/2016/03/13/sunday-bayes-a-brief-history-of-bayesian-stats/>.
- [16] Dhange, N.R., Patil, J., Dhakate, J.P.M., Hingnekar, H. (2023). Literature review for study of characteristics of traffic flow. *International Journal for Research in Applied Science & Engineering Technology*, 11(V): 2393-2405. <https://doi.org/10.22214/ijraset.2023.52079>
- [17] Matowicki, M., Pribyl, O. (2023). On quantification of traffic congestion impacts on socio-economic aspects in cities. In *2023 Smart City Symposium Prague (SCSP)*, Prague, Czech Republic, pp. 1-6. <https://doi.org/10.1109/SCSP58044.2023.10146238>
- [18] Rahman, M.M., Joarder, M.M.A., Nower, N. (2022). A comprehensive systematic literature review on traffic flow prediction (TFP). *Systematic Literature Review and Meta-Analysis Journal*, 3(3): 86-98. <https://doi.org/10.54480/slr.m.v3i3.44>
- [19] Avotins, A., Tetervenoks, O., Adrian, L.R., Severdaks, A. (2021). Traffic intensity adaptive street lighting control. In *IECON 2021-47th Annual Conference of the IEEE Industrial Electronics Society*, Toronto, ON, Canada, pp. 1-6. <https://doi.org/10.1109/IECON48115.2021.9589940>
- [20] Gani, M.H.H., Khalifa, O.O., Gunawan, T.S., Shamsan, E. (2017). Traffic intensity monitoring using multiple object detection with traffic surveillance cameras. In

- 2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), Putrajaya, Malaysia, pp. 1-5. <https://doi.org/10.1109/ICSIMA.2017.8311983>
- [21] Xing, J., Wu, W., Cheng, Q., Liu, R. (2022). Traffic state estimation of urban road networks by multi-source data fusion: Review and new insights. *Physica A: Statistical Mechanics and Its Applications*, 595: 127079. <https://doi.org/10.1016/j.physa.2022.127079>
- [22] Vlasov, A.A. (2022). Robust forecasting of traffic flow intensity. *International Journal of Advanced Studies*, 12(2): 7-20. <https://doi.org/10.12731/2227-930X-2022-12-2-7-20>
- [23] Rao, S.G., RamBabu, R., Kumar, B.A., Srinivas, V., Rao, P.V. (2022). Detection of traffic congestion from surveillance videos using machine learning techniques. In 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Dharan, Nepal, pp. 572-579. <https://doi.org/10.1109/I-SMAC55078.2022.9987342>
- [24] Rajalakshmi, M., Khare, A., Patel, P. (2023). Traffic vehicles detection using deep learning techniques. In 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, pp. 1-5. <https://doi.org/10.1109/ICCCI56745.2023.10128429>
- [25] Someswari, P., Cristin, R., Daniya, T. (2023). Traffic sign detection using deep learning techniques. In 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, pp. 1-6. <https://doi.org/10.1109/ICAECT57570.2023.10118165>
- [26] Hou, B., Zhang, K., Zuo, X., Zhao, J., Xi, B. (2022). PIoT malicious traffic detection method based on GAN sample enhancement. *Security and Communication Networks*, 2022(1): 9223412. <https://doi.org/10.1155/2022/9223412>
- [27] Bosso, N., Magelli, M., Zampieri, N. (2022). Monitoring systems for railways freight vehicles. *International Journal of Computational Methods and Experimental Measurements*, 10(4): 359-371. <https://doi.org/10.2495/CMEM-V10-N4-359-371>
- [28] Oladunjoye, S.A., Ayedun, O.Q. (2019). Analysis of traffic pattern and congestion on Ipaja-Ikotun road corridor of Lagos metropolis, Nigeria. *Ilaro Journal of Environmental Research & Development*, 3(1): 167-182.
- [29] McNaughton, D. (2019). Richard price. In the *Stanford Encyclopedia of Philosophy* (Winter 2024 Edition). <https://plato.stanford.edu/entries/richard-price/>.
- [30] Wikipedia. (2019) Bayes' Theorem. https://en.wikipedia.org/wiki/Bayes%27_theorem.
- [31] Srivastava, T. (2023). 12 Important model evaluation metrics for machine learning everyone should know (Updated 2023). <https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/>.