2nd International Conference on Information Science and Technology Innovation (ICoSTEC)
February, 25 2023. Yogyakarta, Indonesia

Smart Security in Smart City Using Naïve Bayes and RSA

Junaidi[1], R. Roslina[2], B.Herawan Hayadi[3]

[1,3]*Magister of Computer Science, Potensi Utama University*
JL. KL. Yos Sudarso Km. 6,5 No. 3-A, Medan

[2]*Departement of Computer and Informatics Technology, Politeknik Negeri Medan*
JL. Almamater No.1, Padang Bulan, Kec. Medan Baru, Medan

[1]junaidy2906@gmail.com
[2]roslinaich@gmail.com
[3]b.herawan.hayadi@gmail.com

***Abstract*— As all major cities adopt the concept of smart cities, concerns arise among the public regarding data security and privacy. The constant threat of attacks on confidentiality, integrity and accessibility of data makes it vulnerable to cyber attacks. The increasing use of IoT devices also increases the potential for cyber attacks that can harm all IoT users. Therefore, it is crucial for city governments to be aware of data security issues related to smart spaces, services, and citizen security, and to provide solutions to existing problems by making maximum policies related to the implementation of smart city concepts. From the above explanation, the author is taking the analysis step with the title "Analysis of Naive Bayes Classifier and Rsa (Rivest Shamir Adleman) Combination in Smart Security in the Implementation of Smart City in Pemko Medan" where the benefits that can be obtained are to gain deeper understanding of Smart Security level, obtain information about the Smart Securty level, and classify the stage of Smart Securty using the combination of Naive Bayes Classifier and Rsa (Rivest Shamir Adleman) in the implementation of Smart City in Pemko Medan.**

***Keywords*— Naïve Bayes Classifier, Rivest Shamir Adleman, Smart Security**

I. INTRODUCTION

The existence of a Smart City in Indonesia is not only beneficial for one party, the government or citizens, but it is a solution for all [1]. Furthermore, Smart City can increase the efficiency and effectiveness of work, thus improving the quality of life for every element of the city [2]. As one of the largest cities in Indonesia, Medan has begun to develop a system for establishing the Medan Smart City—a city that performs at its best, effectively and efficiently, when managing resources [3]. A Smart city cannot exist without computers, the internet, and intranet networks for data transfer; therefore, security is a crucial issue and a must-have for its establishment [4]. This research aims to build smart security capable of automation data security with a combination of machine learning and cryptography.

The Naive Bayes Classifier method is one of the algorithms in classification techniques that use probability and statistics presented by the English scientist Thomas Bayes, which predicts the likelihood of the future based on past experiences, thus known as Bayes' Theorem [5]. Some studies have shown the feasibility of this algorithm in building smart-system, such as the intelligent system for student personality classification [6], the system for NPC braking decisions in a racing game [7], and the new student admission recommendation system [8]. We choose the Naive Bayes algorithm for the machine learning method, using these studies as a reference in building smart security.

The Rivest-Shamir-Adleman (RSA) algorithm is a cryptographic method with a high level of security because it uses a combination of the results of two prime numbers as the key [9]. The security of the RSA algorithm lies in the difficulty of factoring prime numbers in the formation of the key, so as long as the value of the prime number factor is unknown, the data will remain safe [10]. Recent studies have proven the robustness of the RSA algorithm in securing data, such as study A, study B, and study C.

In this research, we choose the combination of Naive Bayes and the RSA algorithms to build a smart-security system based on the data obtained from PEMKO Medan. The Naive Bayes algorithm will function as the data security level determiner for the given dataset. We evaluate the model performance using the 10-fold cross-validation to measure the accuracy, precision, and recall produced [11]. The RSA algorithm uses the security level output to choose the size of the key for the encryption and decryption process. Finally, the combination of both models yields a smart-security system worthy of implementation in a smart-city system [12].

Meanwhile, Indonesian Cloud Forum Advisor Mochammad James Falahuddin said the most crucial security issue for humans is not running applications. “You see the NSA leaked because Edward Snowden didn't talk. There is no doubt that the system was made by the NSA," he said [13]

II. METHODS

A. Naïve Bayes Classifier

In contrast to iterative iteration principles, Naive Bayes applies the Maximum Likelihood principle throughout the training process, which is more effective [14]. The application of Bayes' theorem allows the Naive Bayes model to classify data based

on prior and posterior probability, combined with evidence to form the formula in equation (1) [15].

$$P(H|D) = \frac{P(D|H) * P(H)}{P(D)} \quad (1)$$

The explanation about equation (1): D is the data with an unknown class, H is the Hypothesis on D in a specific class, P(H|D) is the posterior probability (Probability of H based on condition D), P(D|H) is the prior probability (Probability of D based on condition Q), P(H) is the Probability of H, and P(D) is the Probability of D.

In this research, we use Naive Bayes as a classifier in determining the level of data security [16]. When selecting the encryption bit of the data security process, the RSA method uses the classification output as a reference, with the configuration shown in Table I.

TABLE I
CONFIGURATION OF SECURITY LEVEL

Classification Output	RSA Key Size
Level 4	512 bit
Level 3	1024 bit
Level 2	2048 bit
Level 1	3072 bit

The security level configuration in Table I is the base for the RSA algorithm to secure the data. When the output is a Level 1 security, the RSA algorithm will use a 512-bit key to encrypt and decrypt the data, a 1024-bit key for Level 2, a 2048-bit key for Level 3, and a 3072-bit key for Level 4 [17].

B. Rivest-Shamir-Adleman

The size of the key, when used with the RSA algorithm for encryption, defines the output's level of security; the larger the size of the key, the more secure the data [18]. The RSA key generation uses several parameters, such as p (the first prime number), q (the second prime number), n (the product of p and q), e (the public exponent), and d (the secret private exponent) [19]. All these parameters combined produced a formula for RSA encryption and decryption, as shown in equations (2) and (3) [20].

$$Encryption(n) = n^e \, mod(N) \quad (2)$$

$$Decryption(n) = c^d mod(N) \quad (3)$$

We use the RSA algorithm to produce a model for securing the dataset based on the security level generated from the Naive Bayes algorithm.

C. Dataset

We collected the personnel dataset used for this study from PEMKO Medan, consisting of 391 data. Table II shows the sample from the raw data we obtained. Out of these 391 data, we use 305 data as the training data and 87 as the testing data. The training data will be used to train the model to classify the security level, while the testing data for implementing the model.

TABLE II
SAMPLES OF RAW DATA

R	YS	YB	G	J	SL
IV C	2017	1964	L	Expert Staff	1
IV B	2022	1979	L	Secretary	I
III D	2021	1984	P	Vice Director	II
IV A	2020	1984	P	Head of Division	II
III C	2019	1990	L	Secretary	III
IV A	2016	1972	L	Head of Division	III
III C	2018	1979	L	Village Chief	IV
III C	2019	1977	L	Head of Sub-Division	IV

Notes: R = Rank, YS = Year of Service, YB = Year of Birth, G = Gender, J = Jobs, SL = Security Level

We normalized the data to ease the classification process by changing the data value into a numerical value. Table III shows the normalization configuration, and Table IV shows the results.

TABLE III
NORMALIZATION CONFIGURATION

Category	Old Value	New Value
Rank	III A	1
	III B	2
	III C	3
	III D	4
	IV A	5
	IV B	6
	IV C	7
Year of Service	Year Number	2022 – Year Number
Year of Birth	Year Number	2022 – Year Number
Gender	L	1
	P	0
Jobs	Village Chief, Sub District Head	1
	Secretary Assistant, Assistant Inspector	2
	Secretary, Expert Staff, Inspector	3
	Head of Sub-Division, Head of Section, Head of Agency, Head of Unit, Head of Division	4
	Vice Director, Head of Department, CEO, Director	5
Security Level	I, II, III, IV	1, 2, 3, 4

TABLE IV
SAMPLES OF NORMALIZATION

R	YS	YB	G	J	SL
7	5	58	1	3	1
6	0	43	1	3	1
4	1	38	0	5	2
3	3	1	1	3	3
5	6	2	1	4	3
5	4	44	0	3	3
3	4	43	1	1	4
3	3	45	1	4	4

D. Model

First, we use the Naive Bayes algorithm to classify the normalized dataset to produce the security level. Then, we use the security level configuration to determine the bit for the RSA key and secure the dataset using encryption. To decrypt the data in the dataset, the user must first know the security level to determine the bit for the RSA key, thus improving the data security. Figure 1 shows the complete architecture of the smart-security system used in this research.

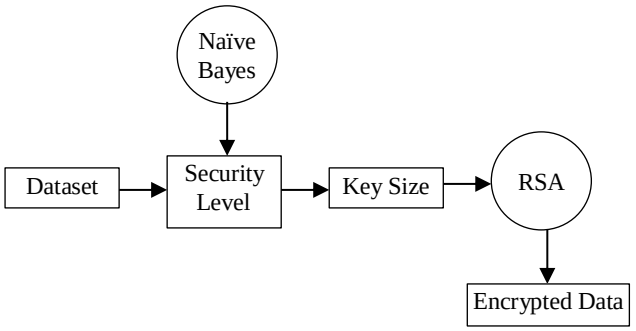

Fig. 1 Model Architecture

For future implementation in the Smart City system, especially in PEMKO Medan, we propose the system architecture as shown in Figure 2.

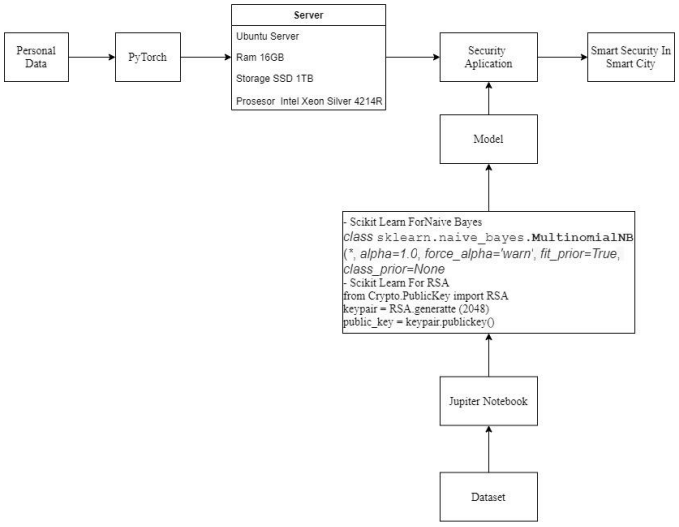

Fig. 2 Proposed System Model

In Figure 2 above, we propose an integrated system between the model and the Smart City system used by PEMKO Medan. We divide the system into two parts, namely the Smart City and Smart Security systems. The Smart City system accesses personal data from PEMKO Medan, which is connected to an Ubuntu server with specifications as shown in Figure 2. The Smart City system accesses personal data from PEMKO Medan, which is connected to an Ubuntu server with specifications as shown in Figure 2. At the same time, the Smart Security system accesses the training dataset to build the model with the help of Jupyter Notebook, with the help of Scikit-learn, utilizing sklearn.naive_bayes and Crypto.PublicKey.RSA libraries. The model built can be implemented using the PyTorch framework, resulting in a Security application. The previous Smart City system is then connected to the Security application to secure personal data when accessed by users.

E. Evaluation

In this study, we evaluate the classification outcomes using the 10-fold and 20-fold cross-validation. We use the accuracy, precision, and recall values to rank the model's performance, using formulas as shown in equations (4) to (6) [21].

$$Accuracy = \frac{TP+TN}{Predicted+Actual} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{Positive\ Predicted} = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{Negative\ Predicted} = \frac{TP}{TP+FN} \quad (6)$$

Using equations (4) to (6), we calculate the accuracy, precision, and recall values from the model's classification result. These values are used to determine whether the model's performance is good enough with a standard of 85%.

III. RESULTS

The first result from the classification process produced a confusion matrix, as shown in Table V. This table displays the prediction result from the model using the Naive Bayes algorithm in classifying the personnel dataset.

TABLE V
TRAINING DATA CLASSIFICATION RESULT

Actual	Positive Prediction			
	1	**2**	**3**	**4**
1	26	3	0	0
2	1	25	9	0
3	0	2	7	5
4	1	0	7	219

We use equations (4) to (6) to analyze the model's performance by analyzing the accuracy, precision, and recall values, as shown in Table VI.

TABLE VI
EVALUATION FOR TRAINING DATA

Accuracy	Precision	Recall
0.908197	0.925538	0.908197

From the result shown in Table VI, the model performance displays over 90% for each accuracy, precision, and recall value. This result shows that the model has a good performance in the training classification and can be used for the next step, which is classifying the testing data for the security level of the RSA encryption.

In the second step, we use the testing data to produce the security-level classification result for the encryption process. Table VII shows the confusion matrix for the classification result of the testing data.

TABLE VII
TESTING DATA CLASSIFICATION RESULT

Actual	Positive Prediction			
	1	2	3	4
1	2	2	0	0
2	1	28	6	0
3	0	0	10	0
4	0	1	2	35

We use equations (4) to (6) to analyze the model's performance by analyzing the accuracy, precision, and recall values, as shown in Table VIII.

TABLE VIII
EVALUATION FOR TESTING DATA

Accuracy	Precision	Recall
0.862	0.895	0.862

From the result shown in Table VIII, the model performance displays over 86% for each accuracy, precision, and recall value. This result shows that the model has a good performance in the testing classification and can be used for the next step, which is the encryption with the RSA algorithm.

In this step, we pick each data corresponding to the security level from the testing data, as shown in Table IX. We choose the first testing data (with Security Level = 3), the second data (Security Level = 1), the sixth data (Security Level = 2), and the tenth data (Security Level = 4).

TABLE IX
DATA FOR RSA ENCRYPTION PROCESS

Data	R	YS	YB	G	J	SL
1st	III D	2021	1985	L	Head of Division	3
2nd	III D	2015	1983	L	Head of Division	1
6th	III D	2021	1979	L	Head of Division	2
10th	IV A	2019	1972	L	Head of Division	4

With the security level from Table IX, we encrypt the data with the corresponding keys, with the result shown in Table X for the 1st data, Table XI for the 2nd data, Table XII for the 6th data, and Table XIII for the 10th data.

TABLE X
ENCRYPTION RESULT FOR 1ST DATA

Data	Encryption Result
R	kchvnHha0V9nPw2CpPSpN5SdhYA76pP+VevzeJIq1AkiFFUY4t8MaKltcLKYLxJLlNZLi8mkZ6jXLzOwIyAZLWCMOQTGAA5nGqpTELEQwWRwaS2PwpfOUp7dN2j3xc1O/ZdJN6teOqXD5dmvOPoHBXFwJqeZm3ESQH+BAxUT/ao=
YS	IdAwxz3V0C7PtfZ8xAkL5pbIkd7bkNQHIa9XG3nfzBFC8rKnbAYW5P1MhjgeSwP2Trl2txfJbWifdgZ0o5dUtDeBz4DKvLNWuGmZ5pvKyvEYtXhQsexVV4oZYENBU/2mEmgHegztsa63IwvdYD2Zw0/Z4maBYHsPxP1vRA6c7o4=
YB	COpsdxlcjj/1SZCfDjz/p1T88Ekzgf52W4k9RApCe94Mw4/zlAhOR1/0S+n6vAGpizifOxiveanPbfK6QbxVGKWU/mpx2a7dJ4pkKgaqgqSnlEeolscxYUoqljsc0wNdaHz+X0jtWzL+D8hTgxyWO/vfz6VYGpKkU7rmkf5NF3k=
G	jvXzHtY9Lt3J0M/sXNsdfac2XR7Uf9v9lA1kgXvACK/Qoz3HE9qN88DOtBr6mNGnSRTW6S2/vhoCkM9I5QqKX4dW5Qi7WMoq6POyC5wibuqlvSEzqpuyVIGINh+4Bjq1ThKSpUZ/oBv0ZJZxkf3j7N2n7x5FZ9cOp98B8FnvzQw=
J	Swx8HTHCVExT7qjHJrLVuYlWB7vaRn3+tXRd1pFKO0PuRzbGkgdQ+sxtdi7XdHvMEJhsU042PeEcKSN5RCO8IkEW8boRfoi/6PhRSPgHXtB9y+7RpbyviZkpdhakEQlAHRxRSKKCqDLcY4O3j8wGMUeZG8czCAjofCXfL24GCnQ=

TABLE XI
ENCRYPTION RESULT FOR 2ND DATA

Data	Encryption Result
R	RtWug2tNWC0n10ZWO+A55LJlI12MhN5nYVqr5sVRTI2dMoCvD7a0uS+z0j10bO+HoG9lAsN2o7eukB8sdq+3ugJHLvPGH+1zV0xbFuEaJq97VHlq++yddG67yneFWTEa3eBxxw6F+NCHJQOZ4TK/YMVPjWDICRBGCSX5Ym+tN5raoEpWQRZxw8SRzEYWjQScQQ4Xo/rBeBi95YqxJCeSeEajkxN1JZmTiUZwSVRYEMDmG33GSAOO8nnq5XKgvL+jx2UStaB4yWeH2b/cBm2TQghFOVr1Gc71kltbZQwOGNlAaZBY9Y0zDGL1JwX4EcxiKFSDDzdPD2CxPPq41aH2qjH5BSmEncOTQJ0IWIqZxJXoF0PXmQKNYdD5ZHzHj2hpA5vz2ToDgfweBd+e+OrapmLosgM/2p/EWiZWkiRFh+5CVsEZDyLtHLqzOokV8zfBaeK//Za9A/PFJJIx5TUbttQmCSWmxuq2XMc5PErFanHXBia4UDdiKyke4i/rbGMO
YS	gwn5IsHtz39nFQB+epng7ZvN43+wELrv4EhI8UcwbHR6J7HM2TYwdolggp4FOhackNmdAtNZRtS8f2i/Nu1OOx3PXCPJ+vQXGCD37NmEIzT+8dkfi0BmJKMzUNozkvixjidpw2EHmAfe2boGGJhsHcHz4Tn/4helxBMNpRcZNP8rmbuz9K9PDAaYUwJIx4tfib/QdRXd9H7sOenocyq9j/EIg//zhsdLUoWpucr12dS2F9tqUkB//CbSzJ0XQPpS+7F3oVPbMRy+5Q2oZ93BxvHPjKh5gE95Fjjl9cX/B5Zy+q1tNW4WNqc3aN4r+lrrwzsNOSV5ErKH+uspgHvZHfHVILEaBKh1W45tmUOSUpYgPHClfRqRn/LmBvGYC6jfULA5TvNwBe57OfmbPdxSEUgwhBZTwgK1K4IhFX8wiFl6pzoN+9gCQs87n4TA844isFuKDAWj9u32EL4jJ4CScvKF2JtubqGkgmUX6lwkxWyztB1v4IgtS7rxKDy28PmE
YB	huiZXhe0Dji2vbVn0Edzxq0JpDhI5XUuWknZ3+EirRvYmLm7/19KhMZ/N/wW32OY79BxLRaJD4JiZK755i2fff6AOjRPBZq4tW0THZvV0NGjPFAqd/Hekz5yigX7igWk5hM23z1DyiLLJQcVjWwUseI1KMVCqBVSTF8XRW72/lHpqtOpttxefTivei5S+/qXInJFb7VbhkGVICoSNDotcgRBIl/oodRmOoxU+KpJ6Sw4mMAsaaiZlS24Nhktnk3tY1HGJXcMUA/5z9jZgJLFtnkbMz33eCX2kKhTOqk1kd8ZjFk4f2C+SM7kS5Ez7a6ILkKxA7jQrqI8ff4sAeeFfjGooziDsM/ZKCYr2BuDoJqArfoGrYkJNjf/mWtqjpJT/zEniBA43ZJTFB4QJrWPR9nXXtf/Pn232AGFC/rnnDhW7ytRIccVpcoIVhf8nApjPXcpAeOGXNtHbnKLG6X5t4+jYddrftnt7Nhk2naVqtU+kEOLb6tzvQ57fs98D/OQ
G	ZyBH7NUiqA7+rQg+b6Y9GH4GNSeuS43FCiq6fUAFE6vP+7v0UhwJH86RTotmdpbjieh0ZjQRDQH3N7qNtI4L1y5YzqI+H4eCWUUoljUHMV3UvW02bBq/hpQLJnlyZUrK0t0d3fREXWqbu3uq8EJE0RrC/XPbHApiFQNNYqVA0vSbQqmgOJUcWmE7u6r2JuximcZ2+VqrffcXDMwmQuQao50N5LPzIruxNVRvw4B3fbxRWmA9F0HVs3IhURUKuXbUJAeVAvncL3WHEJ8TgzD5OJMDk8t6u3BTBVcNmlW+qN6j/lRoqLCClVV0Zeja1XwgLzE8GbBEhSAlIjVCfLQ6iCxckpvXaLT2Et6CQSFWb2J4SsIfZWNUySzWOhdFbnAEjW7mavqaXyGe5+6cSbN7iisLXap8YcwyEhX6NgrU2ErVRs/hvHojM0V+K4lT68dr50fTvdl3JvHf6gO59PHOxlxaFsx0O3ji9qQY8zq/jNDP+KohnZjwFQstBGaCLRjV
J	aaBgOocsBVAJOA27YMc1l1BH+m3/QH2KUF8GfQ2orsIfiAr0MzIRZibvCImnpsW8RD2NApVD4Onuq48w4zDNI8DHNVMIy0KypfKOiyFeGHSPQbU/FDYSxYA5l+9UNerpdt04/JpisWkiou02Yi/nC4P1/TA0q/tzGhvEfrakiXY2XuBroTF2vETTNds9wG6LAd3ch/eWRYUgA7AeJRnl4DVBZR8XFJs1Onlxp0PwiMex7RprdtHkKiXtfOPvmd5Y0jVaprPeT7cU6FnKtMD7W0w8Mc/WMwD

	uq3tc4OhVFZGDVcy6dBA69lSt8/ppFUK8AcZPxd2MRxqq3Q5+FZlHFkPQYSbGBRQAFhv9JdV/ffTJAJPNRf3RLq+gfwMsotPEsyh4OT50K8vu2YiR/m4TtAyaXnEIk/F3co9jeYH1Zn4VFBKRNOwesTs4iS2bx1Zudz9xJYd7kSkBJu7BTN2D5XZi67dNb1AiVf2efyGwbGRFx93KZU44JkvW1nNiRbsN

TABLE XII
ENCRYPTION RESULT FOR 6TH DATA

Data	Encryption Result
R	ARA5pTTuNncva7QSfNnst4/WN+NPKnONUiotm9yOyTl5bme586ErV/C2pPV9AbTCClTVWmTKm98dtS6zTBVBNrvORq6ARSblQumARVf6va5OfNNO3yeM2Is1kbkROE8y6H+f0BmRNcsoGJ5b/ySVFQvA5gAgFAngfodQvCnpUc9eEB8kPBBlFnXHoamt0vDCn3/aUPETbVljmXPfHLQLnBtN+0lEV8wjwqx0fHqzSGkvPpqdjhWHrD/tqlveOLr4gjuByHTf58gdB51Wkj5xOvmHgeodYkc+jfRoNzoAWddXE7J4ITTwP7tXicRUCi5GUVNk1hvbz5GVj+n262VJEQ==
YS	aNhupyfv6y00wpowT/LmkgezL6CS7eNs6ClP4RKLabbDUG51s4nAvWEOh3I3VrVp7bp87MNORN5dyh/KfsEBlGbwTlmhtdMTiKbYDxVWVmbxN6xJjU5q8W+P+PskxJP5Y2yWPPdKMjhNHNXJ1VDor4Bsb6+xjWPVDnfed2PNvQ67nz2pUFm7JNjDM1zdVcDTuHbJsT+zofHc8QjpGGclZ9w+TjFFxo059oGuHSSUnOtcLbmaYyyZrP7sraFJFQN85YfXZSdlwk1fuKtA2/VpzZx3Q+KSZuJ9eecNZB5ruT6lPMwvPY8eUCYDl9GIo28/1Wb5bVxJ/b5yRz3sHTMgzg==
YB	wEyZLzZYPACPwKaf/uYyZQUAPLn8OWQ/jha4GE/Eu6b7MUw3bNjIB9HmVYUlvIZ5s90y9DLdz2tCHc67RHFEIcW1b8PMXU7Dxxya+TKmhTBzvkenQ6SgLop9zWTjjcFqC+THm4JyJq1MjgFWiYK9P4J5Ju/BjkwLH7E3EIxhiTy+dDr/QvOfOPZd7NlOB9gX3hxYCG4BihqKd9/ALPJNwlgDID7e1tIUX5MjSjjNuyDEpT/j5upxtA2sQo6Fhtrh0TrxkA1an1gp1jJNmLS4SVBeWq1DljMnAljGXLcDrcT6TTT1aYXxGjjyZPyCBVG1ZGT6FGBPUCfpQ1OKSicFw==
G	gK5Re58dHC8RYjyPhfKe1kG2jB5mz0cafn9K4wC+C1Eo9VJqSI1haITkLTbDvnWPPQdMtxwkYm9Vmt+exsN2Q2J1SXu7h6MksYRHWfFW1bUhJl1mdkqnqd1SSj07yzD9LmCcCGx38mJCsV5YCIeZn21fNpSxNUI3tp/vxG2iMkO8IDAYgNz+wGv3G7shxnH5WjcH0P4Rp2sYVpW4DcZn6TaIF1FMnGeppY14E81f57jdnemtkV9GVs5k5nzQIJuN8ln5HxzPMLFuNzb6naPXnXJOfS6h+g33VhAWIG+83HJhnRMOG0FSl4bvp4qfiE3hDZcs+ryu4dP/YCR6d9i7rQ==
J	NbmHp3zvms6WRnRuwMrFo0w2RRTA2bOq3GXJwZn0JYz9IgI3D9DeBJDTR+FWKhW45GOgLf45TFNnahuBZXsVQ0Mi2gj2jK9JNsAKAM9ltf/xr/BvS9fuquy9chsGKpg4jiIZoGxxFpSHb8IhT8E041Dy+XLUUtzJvbmjqqfnPQyrIyENBvRPNwL6akx29hTJoZ2ARW8wvjjoOeFgIeqPbpxUhePWN38lj6MOANo2PrBus1kKGeGeIIny6PEltsJNkI9FYJ8rXNZB5HmuN0HtyjSaEHeaZQz5hwZ00Fp66jmqhKljHyfrzuRR9KPJzp/1H2BFxW/URi/DwE5RnMZ97g==

TABLE XIII
ENCRYPTION RESULT FOR 10TH DATA

Data	Encryption Result
R	KOj94xtQX3cxNx2XKXxzWXLKAuFBG1m68O79w/UbXk/KRn11L5lAaTOD2kG1zJ3u0H63osenQZs0PC4isGyL/Q==
YS	bCGRNY6H8/aFSu5c7M0P1DMIJS0oNUaB2dvwpABGmJIf9KLA6Kd02BTMHbC1Bt21fN6rnl6uiERgzVcUkoYXyA==
YB	QIRzd0CSTTKrvjbXuwEwKtfY7OCmlPV52/aqbxVwI9n43urqYG7CZfoS90DUW3byY6h8qCdtuksdewEfGCAH7w==
G	c3y41Fnt/8DRYN5m9mNWQDh1CwI/2oEQi8chkjdmkAaKxW2gMH50wvGXUJ+Ald647ynutrs2ixI8wb3k1R3Psg==
J	mDgRpQEVpBy8BAAOjmPnDJtuAkn6/2Y1o2sM43ryF5zG4HCa6deVycwiIXUA3KgyXuB9b0XgnC7+cKMzUIbECg==

From the result shown in Table X to Table XIII, we found that the RSA model managed to secure the data correctly using the security level from Table IX. Comparing each category in the dataset, we found the result produce different encryption results, especially in terms of the length.

IV. CONCLUSIONS

The results from this research show that using a machine learning method, especially the Naive Bayes algorithm, its feasible to build a smart-security system. By classifying the data in the dataset, we can create a security level to determine the importance of each data. This security level allows the cryptography algorithm, in this case, the RSA algorithm, to encrypt each data with variants of protection level. The RSA algorithm encryption results display the difference in cipher text produced using the security level. One advantage of employing this method is that by using more than one key size, the protection level of the data will adjust according to the data's significance. From the classification result, we conclude that the Naive Bayes algorithm is compatible with building the smart-security system. The accuracy, precision, and recall values show that all values exceed 90%. The combination of the Naive Bayes and the RSA algorithms shows the feasibility of creating an intelligent and automatic system for protecting the data stored in a smart city server.

REFERENCES

[1] I. Widiyastuti, ST., MT, D. Nupikso, N. A. Putra, and V. A. Intanny, "SMART SUSTAINABLE CITY FRAMEWORK: THE SUSTAINABLE AND INTEGRATIVE SMART CITY PROPOSED MODEL," *J. PIKOM (Penelitian Komun. dan Pembangunan)*, vol. 22, no. 1, p. 13, 2021

[2] M. Iqbal, "Smart City in Practice: Learn from Taipei City," *J. Gov. Public Policy*, vol. 8, no. 1, pp. 50–59, 2021

[3] A. Suhendra and A. H. Ginting, "Local Government Policy in Building Smart City in Medan City," *Matra Pembaruan*, vol. 2, no. 3, pp. 185–195, 2018

[4] V. A. Molchanova, "Smart city in the global system of developmental governance," *E3S Web Conf.*, vol. 224, no. December 2016, 2020

[5] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 53, 2017

[6] D. Fahrudy *et al.*, "Intelligent System for Classification of Student Personality," vol. 5, no. 1, pp. 1–9, 2022

[7] S. W. Sanjaya, A. Muhammad Aminul, and T. Afirianto, "Application of Naïve Bayes for NPC Braking Decision in Racing Game," *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*, vol. 3, no. 4, pp. 3252–3257, 2019.

[8] A. Suryadi and E. Harahap, "New Student Admission Recommendation System Using Naive Bayes Classifier at the Indonesian Institute of Education," *Joutica*, vol. 3, no. 2, p. 171, 2018,

[9] M. Jannah, B. Surarso, and Sutimin, "A combination of Rivest Shamir Adlemann (RSA) and Affine Cipher method on improvement of the effectiveness and security of text message," *J. Phys. Conf. Ser.*, vol. 1217, no. 1, 2019

[10] I. A. Achmad Wahyu Hidayat, Riza Arifudin, "Implementation of RSA and RSA-CRT Algorithms for Comparison of Encryption and Decryption

Time in Android-based Instant Message Applications," *J. Adv. Inf. Syst. Technol.*, vol. 2, no. 2, pp. 1–10, 2020, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/jaist/article/view/44302.
[11] M. Ibtasam, "Accuracy Measurements and Decision Making by NaÃ¯ve Bayes and Forward Chaining Method to Identify the Malnutrition Causes and Symptoms," *Sci. J. Informatics*, vol. 8, no. 2, pp. 320–324, 2021
[12] R. Darari, E. Winarko, and A. Damayanti, "Encryption and Decryption Application on Images with Hybrid Algorithm Vigenere and RSA," *Contemp. Math. Appl.*, vol. 2, no. 2, p. 109, 2020
[13] D. Ayu Suci Ilhami, "Data Privacy and Cybersecurity in Smart-City: A Literature Review," *J. Sains, Nalar, dan Apl. Teknol. Inf.*, vol. 2, pp. 2807–5935, 2022, [Online]. Available: https://journal.uii.ac.id/jurnalsnati/article/view/23908/14153.
[14] Heliyanti Susana, "Application of Naive Bayes Classification Method Model on Internet Access Usage," *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022
[15] K. L. Kohsasih and Z. Situmorang, "Analisis Perbandingan Algoritma C4.5 Dan Naïve Bayes Dalam Memprediksi Penyakit Cerebrovascular," *J. Inform.*, vol. 9, no. 1, pp. 13–17, 2022.
[16] R. Firmansyah, E. Utami, and E. Pramono, "Evaluation of Naive Bayes, Random Forest and Stochastic Gradient Boosting Algorithm on DDoS Attack Detection," pp. 1–6, 2022.
[17] R. Abid *et al.*, "An optimised homomorphic CRT-RSA algorithm for secure and efficient communication," *Pers. Ubiquitous Comput.*, 2021
[18] H. T. Sihotang, S. Efendi, E. M. Zamzami, and H. Mawengkang, "Design and Implementation of Rivest Shamir Adleman's (RSA) Cryptography Algorithm in Text File Data Security," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020
[19] K. Suresh, R. Pal, and S. R. Balasundaram, "Two-factor-based RSA key generation from fingerprint biometrics and password for secure communication," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3247–3261, 2022
[20] J. P. Sikumbang, "Combination of Shamir Adleman's Elgamal Rivest Algorithm and Least Significant Bit in Securing Files," *Infokum*, vol. 10, no. 1, pp. 370–379, 2021, [Online]. Available: http://seaninstitute.org/infor/index.php/infokum/article/view/317%0Ahttp://seaninstitute.org/infor/index.php/infokum/article/download/317/251.
[21] D. Pardede, I. Firmansyah, M. Handayani, M. Riandini, and R. Rosnelly, "Comparison Of Multilayer Perceptron's Activation And Optimization Functions In Classification Of Covid-19 Patients," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 8, no. 3, pp. 271–278, Aug. 2022.