# IMPLEMENTATION ALGORITMA NAIVE BAYES USING IDENTIFICATION HUMAN SEX OF DENTAL CANNINUS PANORAMIC

**Nur Nafi'iyah[1], Siti Mujilahwati[2]**

**[1] Informatic Enginering University of Islamic Lamongan
Lamongan, Indonesia
mynaff26@gmail.com**

**[2] Informatic Enginering University of Islamic Lamongan
Lamongan, Indonesia
moedjee @gmail.com**

## *Abstract*

*In this study the introduction to humans automatically using a computer is necessary. Due to the occurrence of accidents or fires that result in difficult identification. It could also be because of a crime, which resulted in need of information who the perpetrators of the crime. So researchers make a system for human identification using dental radiographs. Based on the case of plane crash, fire, Bali Bombing, Stunami in Aceh. So it takes an automated system using a computer in identifying humans to find out the identity of the victim. In order to know the success rate of human sex identification system using naive bayes algorithm. The purpose of this study, namely: Produce a new system in identifying human sex based on panoramic images of teeth using naive bayes algorithm. The results of this study, namely the process of identification of human gender using naïve bayes algorithm accuracy values ± 81%.*

**Keywords: *naïve bayes, dental canninus panoramic, identification human sex***

## Introduction

In Nur Nafi'iyah research, Retno Wardhani, 2016 that the process of identification of human gender using dental panoramic photos successfully implemented. The results of the implementation of the application for the identification of the human gender with an accuracy of 80%.

According to Putra, 2009 in some cases, such as aircraft or fire accidents, other biometric characteristics such as faces, or fingerprints can be damaged and can not possibly be identified by such methods. In such cases, the tooth is the right biometrics to use because bones and teeth have the ability to withstand the effects of force and temperature (average up to 1000oC). Dental biometrics have been applied in the handling of disaster victims, including the September 11 bombings in America, the Bali bombing, and the tsunami disaster in Aceh. Identification of victims in these cases is necessary because the status of the victim's death has a substantial impact on the various aspects left behind.

Lukman, 2006 mentions what is meant by the identification of forensic dentistry is all applications of related dentistry disciplines in an investigation in obtaining postmortem data, useful for determining the authenticity and identity of victims and perpetrators for the sake of law in a judicial process and enforce truth.

There are several types of identification through teeth and an oral cavity that can be applied in the application of all related dentistry disciplines to the investigation in the public interest and the judiciary and in making expert certificates according to Lukman, 2006.

If a dentist with a letter of request as a member of the investigator, a member of the identification team, and as an expert witness if the judge is difficult to decide a case in a trial while on the victim's body there are patterns of bite marks, using dentures, and all dental data that has been done from all dentistry disciplines then the judge will ask an expert to ensure the above in order to establish the decision he will take (Lukman, 2006).

In addition to internal and external examinations, the differences between men and women can be seen from the existing bones. One of the identifiable bones to distinguish the sex is the jawbone (Lukman, 2006).
The purpose of this study, namely to apply the Naïve Bayes algorithm in identifying the human sex from a panoramic canine photograph.

Nur Nafi'iyah research, Retno Wardhani, 2016 human sex identification process based on the dental panoramic photo using the backpropagation algorithm. While this research is a continuation of previous research,

where the continuation of the comparison of two algorithms in identifying human gender. Previous research on human gender identification system was built using Microsoft Visual Studio.Net tools and using backpropagation algorithm. In this study, the system will be built using the same tools but the algorithm used, namely Naive Bayes. Of the two systems built, will be compared the level of accuracy, accuracy, and speed. It is hoped that from this research will result in a better system for identifying human gender.

Another algorithm has ever been applied in identifying the sex of humans, namely C4.5. In Nur Nafi'iyah research, Retno Wardhani, 2018 Fuzzy C4.5 algorithm is used to identify human gender. Where the fuzzy algorithm is used to normalize and categorize the centroid value of the canine teeth, and C45 is used for the classification of the sexes. The result of Nur Nafi'iyah research, Retno Wardhani, 2018 in classifying human gender using the fuzzy C4.5 algorithm, that is 72,5%.

The Fuzzy algorithm is also used for premolar and moral classification on periapical images in Martin L. Tangel, Chastine Fatichah, Fei Yan, Janet P. Betancourt, M. Rahmat Widyanto, Fangyan Dong, Kaoru Hirota, 2013.

In the research of Chastine Fatichah, Martin L. Tangel, Fei Yan, Janet P. Betancourt, M. Rahmat Widyanto, Fangyan Dong, Kaoru Hirota, 2015 to diagnose acute leukemia disease using a decision tree fuzzy method. Where fuzzy is used to process the input, so its value is only scaled 0-1. The fuzzy result is used as the decision tree dataset. The decision tree process for learning, which will produce the tree and rule. The Rule is used in diagnosing leukemia.

A forensic radiology is part of a forensic medicine that studies the identification of humans using postmortem radiographic images of different body parts including skeletons, skulls, and teeth. Identification is done by comparing postmortem image (PM) with antemortem recording (AM).

According to P. Bhargavi, S. Jyothi, 2009 that Naïve Bayes classifier can work quickly and incrementally so as to handle discrete and sustainable attributes. Naïve Bayes also has excellent performance in handling real-life data and can make excellent decisions.

Naïve Bayes is a guided classification algorithm. Where to do the classification should be provided data that has been defined fro each attribute or criteria and class variables. To do the classification is calculated based on the probability value of each class against the variables like equation 1.

$$p(H \mid E) = \frac{p(H) x p(H \mid E)}{p(E)} \ldots\ldots(1)$$

Where p (H | E) = is the probability value of the hypothesis against evidence, p (H) = is the probability value of a hypothesis, p (E | H) = probability value of evidence to a hypothesis, and p (E) = probability value of evidence.

**Research methods**
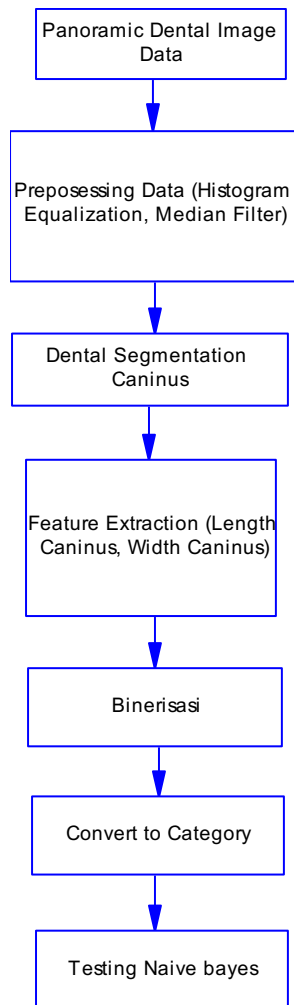
The flow of this study, in Figure 1.

```
┌─────────────────────────┐
│  Panoramic Dental Image │
│          Data           │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│ Preposessing Data (Histogram │
│ Equalization, Median Filter) │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│   Dental Segmentation   │
│         Caninus         │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│ Feature Extraction (Length │
│  Caninus, Width Caninus) │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│        Binerisasi       │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│    Convert to Category  │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│    Testing Naive bayes  │
└─────────────────────────┘
```

**Figure 1.** Diagram Flow Research

Stages in the identification of the sex of humans using the Naïve Bayes algorithm, namely: make image file panoramic gear, prepossessing data. Data prepossessing steps include: equalization histogram, and median filter. Where data in the form of file is done histogram equalization then filtered filter using Matlab. The next step is to segment the image of a panoramic tooth. The resultant image repair data is taken by canine wear using Photoshop.

Segmentation results in the form of canine panoramic image and thresholding. The thresholding process changes the image from gray level to binary form. The next step, perform feature extraction. Feature extraction, taking the sheet and the length of the panoramic canine gear.

Features obtained from the feature extraction stages are numeric values, then scaled to categories. Then test.

Figure 1 is in detail the step in the process of dental panoramic image input, image enhancement (histogram equalization, and median filter), segmentation or cropping (taking upper left canine, lower left canine, right upper right canine and lower right canine teeth) the results of segmentation and then performed binary (black and white), binary results extracted feature width and length of panoramic canine image image, feature data is scaled to categories, and testing.

**Research Results and Discussion**

Data used in this research are 29 panoramic images of teeth. The data areas in Table 1. Data 29 panoramic images obtained from Ibnu Sina Gresik Hospital (RSUD Gresik), the data is taken by recording the teeth of each person to be investigated as many as 29 people using radiology equipment.

**Table 1.** Data of Dental Panoramic Image

| No | Data Image Dental Panoramic |
|----|------------------------------|
| 1 |  |
| 2 |  |
| 3 |  |

From the data in Table 1 then done image improvement, image improvement aims to create a better image. Image improvement process includes histogram equalization and median filter. The results are in Table 2.

The result of the image improvement in Table 2 is then cropped taking the canines of each image. Each image is taken an upper left, right top right, bottom left and bottom right teeth, a total of each image is 4 canine teeth. If the data used is 29 panoramic tooth image then total canine teeth used in this research are 116 images of the canine tooth. The cut of the canines is as in Table 3.
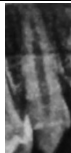
**Table 2.** Result Enhacemant Image

| No | Data Image Dental Panoramic |
|----|------------------------------|
| 1 |  |
| 2 |  |
| 3 |  |

**Table 3.** Pieces of Canine Gear Panoramic Image

| No | Canine | Image |
|----|--------|-------|
| 1 | Left top-1 |  |

| 2 | Left bottom-1 |  |
| 3 | Right up -1 |  |
| 4 | Right bottom-1 |  |
| 5 | Left top-2 |  |
| 6 | Left bottom-2 |  |
| 7 | Right up -2 |  |
| 8 | Right bottom-2 |  |
| 9 | Left top-3 |  |
| 10 | Left bottom-3 |  |

| 11 | Right up -3 | |
|----|------------|---|
| 12 | Right bottom-3 | |
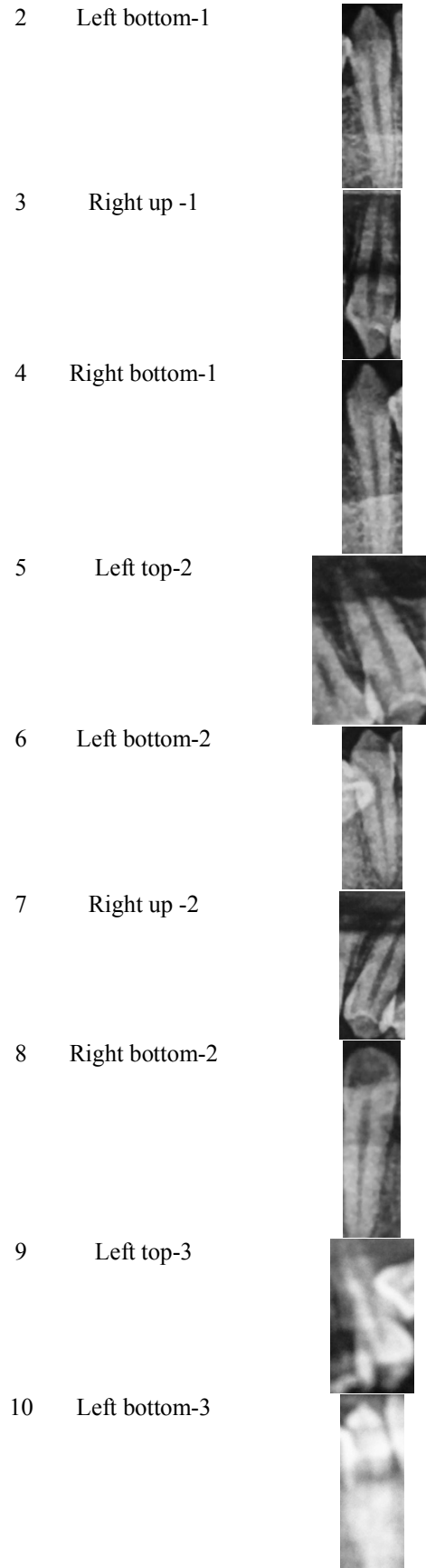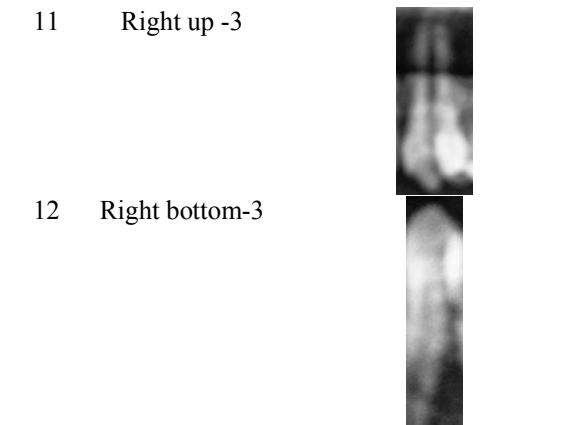
The feature extraction stage is the step to take the value of the features used in the research. Features took from the image of the panoramic canine gear, ie length, and width. So there are 2 variables, ie x is the width and y is the length. After doing the steps in image processing above, then do binary. Image binarization, ie change the image from grayscale to binary form.

Input image in this research is a dental panoramic image in the form of RGB, then do image repair and change to grayscale form. And the next stage makes the form of binary (black and white). The binary process uses an iterative adaptive thresholding algorithm. The binary results in Table 4.

**Table 4.** Result Binerisasi

| No | Canine | Image |
|----|--------|-------|
| 1 | Left top | |
| | Left bottom | |
| | Right up | |
| | Right bottom | |
| 2 | Left top | |

Left bottom

Right up

Right bottom
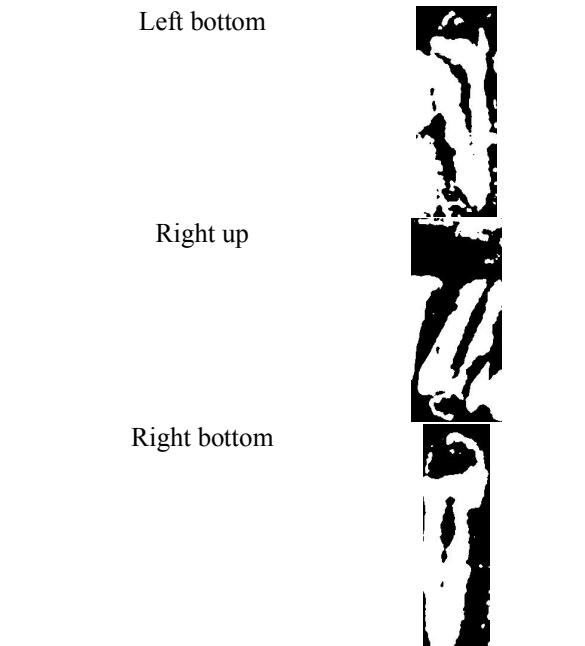
The binary results of Table 4 were then extracted by Matlab. The results are like Table 5. The features taken are the width of the panoramic canine image as the variable x, and the length of the panoramic canine image as the y variable. and sex labeling is done. Where 0 is female gender and 1 is male gender.

**Table 5.** Results of Panoramic Caninus Image Feature

| No | x | y | type |
|----|-----|-----|------|
| 1 | 44 | 150 | 0 |
| 2 | 47 | 150 | 0 |
| 3 | 57 | 176 | 0 |
| 4 | 29 | 118 | 0 |
| 5 | 63 | 173 | 1 |
| 6 | 62 | 200 | 1 |
| 7 | 60 | 160 | 1 |
| 8 | 56 | 195 | 1 |
| 9 | 100 | 193 | 1 |

From Table 5 data in the form of numbers, whereas to conduct the process of testing Naïve Bayes data required data in the form of categories. Data in the form of categories, ie examples: small, large, short, long, and others. So the data used in the Naïve Bayes process must first be changed to the category form. How to make changes to category form, by searching the average value of each variable. For example: variable x (the width of the panoramic canine tooth image is calculated the average value is 60.68), and the y variable (the length of the image of the panoramic canine canal is calculated the average value is 172,71). Then to convert to category form, ie eg data input variable x <= 60,68 then variable x is "Narrow", otherwise variable x is "Width". Next variable y is also so to change to form category, that is example data of variable input y <= 172,71 then variable y is "Short", otherwise variabel y is "Length". Results from changes to categories form as Table 6.

**Table 6.** Result Data Category

| No | x | y | jenis |
|----|--------|-------|-------|
| 1 | Narrow | Short | Woman |
| 2 | Narrow | Short | Woman |
| 3 | Narrow | Long | Woman |
| 4 | Narrow | Short | Woman |

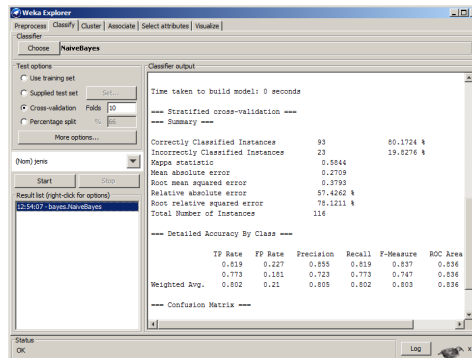| 5 | Narrow | Short | Woman |
|----|--------|-------|-------|
| 6 | Narrow | Short | Man |
| 7 | Narrow | Long | Man |
| 8 | Wide | Long | Man |
| 9 | Wide | Long | Man |
| 10 | Wide | Long | Man |



**Figure 2.** Accuration Result Naïve Bayes in Weka

Where in the process of probability calculation p (E | H) where the value of evidence associated with class Hypothesis 0, it will produce 0. In order not to produce 0. Then in the calculation process there is a modification into Equation 2.

$$p(E \mid H) = \frac{jumlah\_E\_yang\_berkai\tan\_dengan\_H + 1}{jumlah\_keseluruhan\_E + var iabel\_E} \quad ....(2)$$
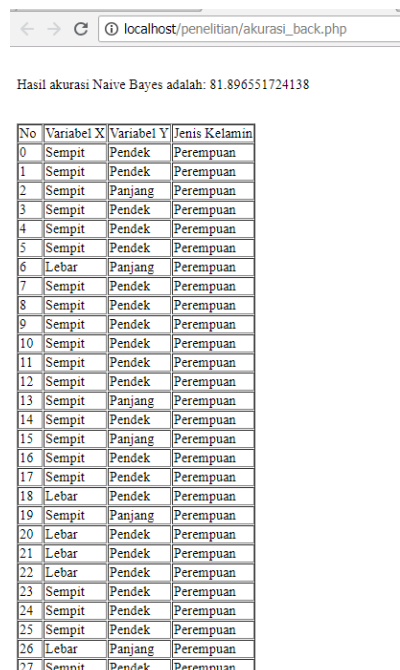


**Figure 3.** Accuration Result Naïve Bayes in Webserver

As for calculating the accuracy value of the Naïve Bayes algorithm the researcher uses 3 tools, that is Web server with PHP language, Weka and Visual Studio Net. The view of the accuracy value of Naïve Bayes with Weka as Fig 2 with an accuracy of 80.17%. The view of the accuracy value of Naïve Bayes PHP language as Figure 3, with an accuracy value of 81.89%. The view of the accuracy value of Naïve Bayes with Visual Studio Net as Figure 4 with an accuracy value of 81.89%.

**Figure 4.** Accurate Result Naïve Bayes in VB Net

## Conclusions and recommendations

From this research can be concluded that the process of identification of human gender with a dental panoramic image using The Naive Bayes algorithm of 81% accuracy value. While in Nur Nafi'iyah research, Retno Wardhani, 2018 gender classification process using The Fuzzy C4.5 algorithm of accuracy value 72,5%. The Naïve Bayes algorithm is of higher accuracy than Fuzzy C4.5.

Suggestions for further research data used for identification should be in greater number, and expected the data can represent the population in Lamongan.

## References

Bilqis Amaliah, Anny Yuniarti, Anindita Sigit Nugroho, Agus Zainal Arifin. 2011. Pemisahan Gigi pada Dental Panoramic Radiograph dengan Menggunakan Integral Projection yang Dimodifikasi. *Jurnal Ilmiah Kursor, Volume 6 Nomor 2 Juli*, 123-130.

Chastine Fatichah, Martin L. Tangel, Fei Yan, Janet P. Betancourt, M. Rahmat Widyanto, Fangyan Dong, Kaoru Hirota. 2015. Fuzzy Feature Representation for White Blood Cell Differential Counting in Acute Leukemia Diagnosis. *International Journal of Control Automation and Systems*, 1-11.

Lukman, D. 2006. *Buku Ajar Ilmu Kedokteran Gigi Forensik Jilid 2.* Jakarta: CV. Sagung Seto.

Martin L. Tangel, Chastine Fatichah, Fei Yan, Janet P. Betancourt, M. Rahmat Widyanto, Fangyan Dong, Kaoru Hirota. 2013. Dental Classification for Periapical Radiograph Based on Multiple Fuzzy Attribute. *IEEE*, 304-309.

Nafi'iyah, N. 2015. Algoritma Kohonen dalam Mengubah Citra Graylevel menjadi Citra Biner. *JITIKA Vol. 9, No.2*, 49-55.

Nafi'iyah, N. 2016. Perbandingan Modus, Median, K-Standar Deviasi, Iterative, Mean dan Otsu dalam Thresholding. *Spirit Vol.8, No.2*, 31-36.

Nur Nafi'iyah, Chastine Fatichah. 2017. Fuzzy Self Organizing Map untuk Proses Thresholding pada Citra Dental Panoramic. *Senasif.* Malang: ITN.

Nur Nafi'iyah, Retno Wardhani. 2016. Sistem Identifikasi Jenis Kelamin Manusia Berdasarkan Foto Panoramik Gigi. *Seminar Hasil Penelitian dan Pengabdian Masyarakat* (pp. 120-125). Jember: Politeknik Negeri Jember.

Nur Nafi'iyah, Retno Wardhani. 2017. Perbandingan Otsu dan Iterative Adaptive Thresholding dalam Binerisasi Gigi Kaninus Foto Panoramik. *JITIKA Vol. 11, No. 1*, 21-28.

Nur Nafi'iyah, Retno Wardhani. 2018. ANALISA FUZZY C45 DALAM MENGKLASIFIKASI JENIS KELAMIN MANUSIA DARI FITUR CITRA PANORAMIK GIGI KANINUS. *Seminar Nasional Inovasi dan Aplikasi Teknologi di Industri* (pp. 160-166). Malang: Institut Teknologi Nasional.

P. Bhargavi, S. Jyothi. 2009. Applying Naive Bayes Data Mining Techniques for Classification of Agricultural Land Soils . *International Journal of Computer Science and Network Security Vol.9, No.8*, 117-122.

Putra, D. 2009. *Sistem Biometrik.* Yogyakarta: Andi.

Sugiartha, I Gusti Rai Agung. 2016. Ekstraksi Warna, Tekstur dan Bentuk untuk Image Retrieval. . *Seminar Nasional Teknologi Informasi dan Multimedia* (pp. 4.3-1-4.3-6). Yogyakarta: STMIK AMIKOM.

Wawan Gunawan, Agus Zainal Arifin. 2017. Lokal Fuzzy Thresholding Berdasarkan Pengukuran Fuzzy Similarity pada Interaktif Segmentasi Citra Panoramik Gigi. *Jurnal Infotel, Volume 9 Nomor 1*, 40-47.

Yuita Arum Sari, Ratih Kartika Dewi, Chastine Fatichah. 2014. Seleksi Fitur Menggunakan Ekstraksi Fitur Bentuk, Warna dan Tekstur dalam Sistem Temu Kembali Citra Daun. *JUTI Volume 12 No 1 Januari*, 1-8.