



A Bayesian Method for Gastric Disease Diagnosis Estimation Among USU Mathematics Students

Masdalipah Harahap^{1*}

^{1*}Departemen of Mathematics, Universitas Sumatera Utara, Medan, Indonesia
masdalipahharahap01@email.com

ABSTRACT

Gastric diseases are frequently experienced by university students, especially Mathematics students at Universitas Sumatera Utara, due to demanding academic activities, irregular dietary patterns, and academic pressure. These conditions often lead to symptoms such as epigastric discomfort, nausea, vomiting, abdominal bloating, and early satiety. Since similar symptoms may arise from different gastric disorders, including dyspepsia, gastroesophageal reflux disease (GERD), gastritis, and Helicobacter pylori infection, an analytical approach is required to estimate the most likely diagnosis based on observed symptoms. This study employs Bayes' Theorem to estimate gastric disease diagnoses using symptom data obtained from questionnaires distributed to students. Prior probabilities were determined based on the proportion of each disease within the observed sample, while conditional probabilities were calculated from the frequency of symptoms associated with each disease category. The analysis was carried out through manual calculations and supported by Microsoft Excel. The findings reveal that GERD exhibits the highest posterior probability in a specific test case, reaching 95%, indicating it as the most probable diagnosis. These results suggest that Bayes' Theorem can be effectively utilized as a supporting method for estimating early-stage gastric disease diagnoses among university students.

Keywords: Bayes' Theorem; Gastric disease; GERD; Posterior probability; University students.

Received: 2 October 2025

Revised: 4 November 2025

Published: 30 December 2025

INTRODUCTION

Gastric diseases are common health problems experienced by university students due to intensive academic activities, irregular eating habits, and psychological stress. Students often neglect balanced meal schedules because of tight lecture schedules, assignments, and examinations. As a result, symptoms such as epigastric pain, nausea, bloating, vomiting, and early satiety frequently occur and may interfere with daily

academic performance. These symptoms are not specific to a single disease and may indicate various gastric disorders, including dyspepsia, gastroesophageal reflux disease (GERD), gastritis, and *Helicobacter pylori* infection. Therefore, an analytical approach is required to estimate the most probable gastric disease based on the observed symptoms to support early diagnosis among university students.

Several previous studies have applied probabilistic and Bayesian approaches in medical diagnosis to handle uncertainty in symptom-based decision making. Bayesian methods have been widely used to estimate disease probabilities by combining prior information with observed evidence, particularly in cases where symptoms overlap across multiple diseases. Recent research has shown that Bayes' Theorem is effective in medical diagnostic systems, including gastrointestinal disorders, due to its ability to update probabilities based on new information.

This study aims to apply Bayes' Theorem to estimate gastric disease diagnoses among Mathematics students at Universitas Sumatera Utara using symptom data collected through questionnaires. The novelty of this research lies in its focus on a specific academic population and the use of a simple Bayesian framework that can be implemented through manual calculations and Microsoft Excel.

METHODS

The research method used in this study is a quantitative descriptive method with a probabilistic approach. This research was conducted at the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara. The objects of this study were Mathematics students who experienced symptoms related to gastric diseases. Data collection was carried out during the academic year 2025/2026.

The population in this study consisted of Mathematics students at Universitas Sumatera Utara, while the sample was determined using a sampling technique. A total of 41 respondents participated in this study. The data used were primary data obtained directly from respondents through a questionnaire distributed using Google Form.

The data collection technique was carried out using a closed questionnaire containing a list of gastric disease symptoms and types of gastric diseases. Respondents were asked to select the symptoms they experienced and the type of gastric disease they had experienced. The variables in this study consisted of symptom variables as independent variables and types of gastric diseases as dependent variables.

The data analysis technique used in this study was the Bayes Theorem method. The analysis began with calculating the prior probability based on the proportion of each type of gastric disease found in the data. Next, the likelihood probability was calculated based on the frequency of each symptom appearing in each disease category. These values were then used to calculate the posterior probability to estimate the most likely type of gastric disease experienced by the respondent.

Data processing and calculations were carried out manually and with the assistance of Microsoft Excel to ensure accuracy and efficiency in probability calculations. The results of the analysis were presented in the form of tables and probability values that show the level of confidence for each type of gastric disease.

RESULT AND DISCUSSION

This section presents the research results and discussion based on processed data obtained from questionnaires that were analyzed using Bayes Theorem. The discussion emphasizes analytical interpretation to identify the significance and contribution of the probabilistic approach in estimating gastric disease diagnoses among university students.

The processed data indicate that GERD has the highest prior probability compared to other types of gastric diseases. This finding suggests that GERD is the most dominant gastric disorder among Mathematics students at Universitas Sumatera Utara. The dominance of GERD reflects common student lifestyle patterns, such as irregular eating schedules, high academic workload, and stress, which are known to increase the risk of acid reflux-related disorders. This result highlights the relevance of using probabilistic analysis to capture disease tendencies within a specific population.

The likelihood analysis demonstrates that several symptoms, including epigastric pain, nausea, vomiting, decreased appetite, bitter or acidic taste in the mouth, and frequent belching, show strong probabilistic associations with specific gastric diseases. The processed prior and likelihood probability values for each gastric disease and symptom are presented in Table 1.

Table 1. Probability Values Of Diseases And Symptoms

Percentage/Probability	Disease Hypothesis			
	X1	X2	X3	X4
P(Ai)	0.317	0.585	0.073	0.024
P(Y1)	0.384	0.75	0.666	0
P(Y2)	0.769	0.791	0.333	1
P(Y3)	0.615	0.375	0	0
P(Y4)	0.461	0.583	0.333	0
P(Y5)	0.23	0.25	0.666	0
P(Y6)	0.384	0,166	0.333	0
P(Y7)	0.076	0.166	0	0
P(Y8)	0.076	0.375	0.333	0
P(Y9)	0.076	0.333	0.666	0
P(Y10)	0.076	0	0	0
P(Y11)	0.153	0.333	0	0
P(Y12)	0.461	0.333	0.333	0
P(Y13)	0.076	0.833	0.333	0

P(Y14)	0.153	0.208	0	0
P(Y15)	0.384	0.208	0.333	0

Rather than evaluating symptoms independently, Bayes Theorem integrates multiple symptoms simultaneously as evidence, enabling clearer differentiation between diseases with overlapping clinical characteristics.

The posterior probability analysis reveals that GERD achieves the highest probability value, reaching approximately 95% in the analyzed case, while other disease categories show substantially lower probability values. The posterior probability results obtained using Bayes Theorem are summarized in Table 2.

Table 2. Bayes Simulation Program Results

	X1	X2	X3	X4
P(Ai)	0.317	0.585	0.073	0.024
P(Y1)	0.384	0.75	0.666	0
P(Y2)	0.769	0.791	0.333	1
P(Y3)	0.615	0.375	0	0
P(Y6)	0.384	0,166	0.333	0
P(Y8)	0.076	0.375	0.333	0
P(Y9)	0.076	0.333	0.666	0
P(B Ai)P(A)	0.000127688	0.0027	0	0
P(Ai B)	0.045191639	0.95481	0	0

This result indicates that the observed combination of symptoms is statistically most consistent with GERD. The high posterior probability confirms the effectiveness of Bayes Theorem in updating prior information using observed evidence to produce a clear diagnostic tendency.

To illustrate these highly consistent results, Figure 1 visualizes the posterior probability of GERD diagnosis, confirming that the observed combination of symptoms has clear diagnostic value.

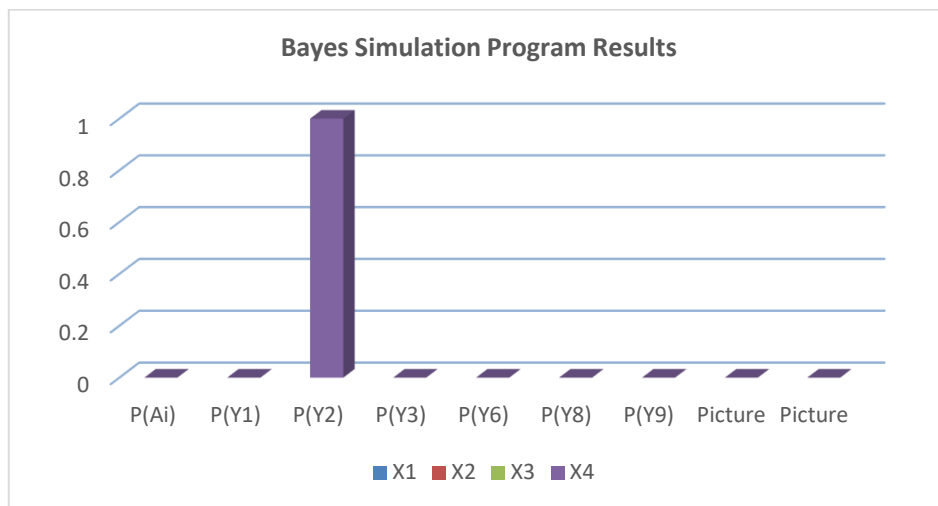


Figure 1. Bayes Simulation Program Result

The significance of this research lies in its contribution to early diagnostic support through a quantitative and probabilistic framework. Unlike descriptive symptom-based approaches, this method provides confidence-based diagnostic estimations that can assist in identifying the most likely disease. The novelty of this study is reflected in the application of Bayes Theorem to real questionnaire-based data from Mathematics students and the use of Microsoft Excel as a practical and accessible analytical tool. These findings demonstrate that probabilistic methods can serve as reliable decision-support tools in the context of gastric disease diagnosis.

CONCLUSION

The study demonstrates that the application of Bayes' Theorem provides an effective framework for estimating gastric disease diagnoses among students of the Mathematics Study Program at Universitas Sumatera Utara. By integrating prior probabilities derived from student symptom data with conditional probabilities for each disease, the model successfully identifies the most likely disease, with GERD showing the highest posterior probability in the examined cases.

This approach confirms that probabilistic methods can support early detection and decision-making in situations with uncertain or overlapping symptoms, offering a structured alternative to self-assessment. The consistency between manual calculations and spreadsheet simulations highlights the reliability and practicality of the method for academic environments.

For further research, it is recommended to expand the model by incorporating larger and more diverse student populations, additional symptom variables, and other probabilistic techniques such as Bayesian networks. Integrating real-time monitoring data or mobile-based symptom tracking could also enhance the accuracy and applicability of early diagnostic estimations in broader healthcare contexts.

BIBLIOGRAPHY

- Afni, N., Salim, A., & Iqbal Maulana, Y. (2021). Sistem Informasi 3,4,5 Fakultas Teknik & Informatika 1,3,4,5 , Fakultas Teknologi Informasi 2 Universitas Bina Sarana Informatika 1,3,4,5. *Universitas Nusa Mandiri*, 5(2). <https://doi.org/10.52362/jisicom.v5i2.659>
- Anggilina, A., & Eviyanti, A. (2021). Web-Based Expert System for Diagnosing Gastric Disease Using Bayes Theorem Method. *Procedia of Engineering and Life Science*, 1(2). <https://doi.org/10.21070/pels.v1i2.944>
- Budi Riyanto, N., & Suria, O. (n.d.). *Sistem Pakar Diagnosa Penyakit Pencernaan Menggunakan Metode Teorema Bayes* 7.
- Getrudis Mali, M., Nababan, D., Herlina Ullu, H., & Baso, B. (2024). Implementasi Teorema Bayes Pada Sistem Pakar Untuk Mendiagnosa Penyakit Asam Lambung Berbasis Website. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 7(1), 182–192. <https://doi.org/10.55338/jikoms.v7i1.2843>
- Homepage, J., Ismaredah, E., & Maulana Sidiq, A. (2025). *MALCOM: Indonesian Journal of Machine Learning and Computer Science Implementation of MERN Stack Technology in an Expert System for Early Diagnosis of Stomach Diseases Using the Naive Bayes Method Implementasi Teknologi MERN Stack pada Sistem Pakar Diagnosis Awal Penyakit Lambung Menggunakan Metode Naive Bayes*. 5, 1559–1569. <https://doi.org/10.57152/malcom.v5i4.2226>
- Kevin arsan kamto, A. siddiq purnomo. (2024). SISTEM DIAGNOSA STUNTING MENGGUNAKAN TEOREMA BAYES. *Jurnal Teknik Informatika Dan Sistem Informasi*, 11(2), 282–294.
- Murni, S., & Riandari, F. (n.d.). *Penerapan Metode Teorema Bayes Pada Sistem Pakar Untuk Mendiagnosa Penyakit Lambung*.
- Nurhaidah, F. S., Anugrah, S. D., Putri, A. F., R Tukloy, W. D., Shofa, K. N., Khairunnisa, S., Primadani, L. H., Wahyudi, T., Kamaruzzaman, A. R., & Nita, Y. (2021). Pengetahuan Mahasiswa Universitas Airlangga Mengenai Dispepsia, Gastritis, dan Gerd beserta Antasida sebagai Pengobatannya. In *Jurnal Farmasi Komunitas* (Vol. 8, Issue 2).
- Polotskaya, K., Muñoz-Valencia, C. S., Rabasa, A., Quesada-Rico, J. A., Orozco-Beltrán, D., & Barber, X. (2024). Bayesian Networks for the Diagnosis and Prognosis of Diseases: A Scoping Review. *Machine Learning and Knowledge Extraction*, 6(2), 1243–1262. <https://doi.org/10.3390/make6020058>
- Purnomo, N., Wardana, B., Yuliana, D., & Rasyid, M. (2025). IDENTIFICATION OF DENTAL AND ORAL DISEASES IN HUMANS USING BAYES THEOREM METHOD. *JOURNAL OF SCIENCE AND SOCIAL RESEARCH*, 8(1), 378. <https://doi.org/10.54314/jssr.v8i1.2591>
- Ramadhan, P. S. (2018). Sistem Pakar Pendeteksian Varicella Simplex Dengan Menggunakan Teorema Bayes. In *JURIKOM* (Vol. 5, Issue 5). <http://ejurnal.stmik-budidarma.ac.id/index.php/jurikom> | Page | 454
- Sagat, N. A., & Purnomo, A. S. (2021). Sistem Pakar Diagnosa Penyakit Mata Menggunakan Metode Teorema Bayes. *Jurnal Pendidikan Dan Teknologi Indonesia*, 1(8), 329–337. <https://doi.org/10.52436/1.jpti.73>
- Sri Winar, Hannan Asrawi, Rekha Audina, & Riyadhul Fajri. (2025). Penerapan Teorema Bayes Pada Sistem Pakar Untuk Mendiagnosa Penyakit Ayam Berbasis Web. *Sisfo: Jurnal Ilmiah Sistem Informasi*, 9(1), 1–9. <https://doi.org/10.29103/sisfo.v9i1.22135>
- Sulardi, N., & Witanti, A. (2020). SISTEM PAKAR UNTUK DIAGNOSIS PENYAKIT ANEMIA MENGGUNAKAN TEOREMA BAYES. *Jurnal Teknik Informatika (Jutif)*, 1(1), 19–24.

<https://doi.org/10.20884/1.jutif.2020.1.1.12>

Yanti, Y., & Sulindawaty, S. (2020). Expert System Diagnose Disease Persian cat with Bayes Theorem Method. *Journal Of Computer Networks, Architecture and High Performance Computing*, 2(1), 149–157. <https://doi.org/10.47709/cnapc.v2i1.380>