

# Application of the Certainty Factor Method to Diagnose Escherichia Coli Bacteria in Refilled Drinking Water

<sup>1</sup>Agustina Simangunsong, <sup>2</sup>Penda Sudarto Hasugian  
<sup>1,2</sup>Teknik Informatika, STMIK Pelita Nusantara, Sumatera Utara  
 Email: [agustina45@gmail.com](mailto:agustina45@gmail.com)<sup>1</sup>, [pendasudarto99@gmail.com](mailto:pendasudarto99@gmail.com)<sup>2</sup>

Keywords	<b>Abstract.</b> An expert system is a software package or computer program package that is intended as a provider of advice and aids in solving problems in certain areas of specialization, such as problems in medicine, education, and so on. The development of expert system development in the field of bacteriology is one thing that is expected to improve the quality of human life. One of them is to apply an expert system using the Certainty Factor method to diagnose escherichia coli (e-coli) bacteria in refilled drinking water. This study aims to increase knowledge about how to make an inference engine to control the process of identifying solutions using the Certainty Factor method and to design an expert system application using the Certainty Factor method to diagnose Escherichia coli bacteria in refill drinking water.
Certainty Factor Expert system Application	

## 1. INTRODUCTION

In line with the increasing demand for drinking water at this time, many depots or places for refilling drinking water have sprung up. Sometimes many parts of the community do not know whether the water is suitable for consumption or not. The reason is that the water that is consumed daily does not go through the cooking or boiling process first [1], [2].

Maybe the drinking water that is bought and consumed looks clean and clear like drinking water in general, but behind the clean and clear drinking water, people cannot know what is contained in it. Cheap and affordable prices may be one of the reasons people choose to refill water from existing drinking water depots. No matter where drinking water comes from, the most important thing is that daily drinking needs are met [3].

Many of the depots that have sprung up today do not include a letter from the local health office stating that the drinking water sold is fit for consumption. Expensive costs and very difficult affairs make drinking water depot owners ignore the most important thing that really must be owned [4], [5].

One of the things that can threaten health through drinking water is the presence of the Escherichia Coli (E-Coli) bacteria. To find out whether drinking water contains E-Coli bacteria is not easy, because it is very small and invisible. One of the consequences that can be caused by E-Coli bacteria is stomach pain, vomiting, diarrhea, high blood pressure, and even kidney problems. From the above problems, it is necessary to apply a method to diagnose E-Coli bacteria in refill drinking water by applying the certainty factor method [6]–[8].

## 2. METHOD

The certainty factor was introduced by Shortliffe Buchanan in making MYCIN (Wesley, 1984). Certainty Factor (CF) is a clinical parameter value given by MYCIN to show the amount of confidence. Certainty Factor is defined as follows [9] :

$$CF(H,E) = MB(H,E) - MD(H,E) \dots \dots \dots (2.1)$$

Information :

CF(H, E) : Certainty factor from hypothesis H which is influenced by symptoms (evidence) E.

Magnitude of CF Ranges from -1 to a value of 1 indicates absolute trust.

MB(H, E) : A measure of belief in the hypothesis H which is influenced by the E symptom.

MD(H, E): A measure of increased disbelief to hypothesis H which is influenced by the E phenomenon.

Certainty factor to determine parallel CF

- CF (x dan y) = Min (CF (x), CF(y)).....(2.2)  
 CF (x atau y) = Max (CF (x), CF (y)).....(2.3)  
 CF (Tidak x) = -CF (x)..... (2.4)  
 Certainty factor to determine CF sequential  
 CF (H,E)1 = CF (E,e) \* CF (H,E)..... (2.5)  
 Certainty factors for similar rules  
 CfcombineCF (H,E)1,2 =CF (H,E)1 + CF (H,E)2 \* (1CF(H,E)1).....(2.6)  
 CfcombineCF (H,E)old3=CF (H,E)old + CF (H,E)3\* (1CF(H,E)old)..... (2.6)

Table 1.  
CF value [10]

Uncertain Term	CF
<i>Definitely not</i> (pasti tidak)	<b>-1.0</b>
<i>Almost Certainly not</i>	<b>-0.8</b>
<i>Probably not</i>	<b>-0.6</b>
<i>Maybe not</i>	<b>-0.4</b>
<i>Unknown</i> (don't know)	<b>-0.2 to 0.2</b>
<i>Maybe</i>	<b>0.4</b>
<i>Probably</i> (possibly true)	<b>0.6</b>
<i>Almost certainty</i>	<b>0.8</b>
<i>Definitely</i>	<b>1.0</b>

### 3. RESULTS AND DISCUSSION

#### 3.1. Diagnostic Analysis of Escherichia Coli Bacteria on Refill Water

The analysis is useful for knowing software requirements and the needs of an expert system that is built. In this stage, a search and collection of data and knowledge required by the expert system are carried out. So that in the end, the results obtained from the analysis are in the form of a system whose structure can be defined well and clearly. The purpose of system analysis in the development of this expert system is to obtain various user and system needs, namely regarding the input and output that must be provided by the user and required by the system. Based on the results of consultations conducted with BTKL (Environmental Health Engineering Agency) Medan city, there are several The characteristics of drinking water contaminated with Escherichia Coli bacteria include:

1. Smells like sewage or mud.
2. Feels bitter or chelate on your tongue.
3. Not clear or cloudy.
4. A little yellowish.
5. Color change in 2 days.

After obtaining the characteristics of drinking water contaminated with escherichia coli bacteria from the BTKL (Environmental Health Engineering Agency), then questions are asked to the experts to get MB (Measure of Belief) and MD (Measure of increased Disbelief) values for each characteristic of drinking water contaminated with Escherichia coli bacteria, such as Table 2 below:

Table 2.  
Question

Feature Code	Question	Answer
C1	Does the refill drinking water	TY TT KY CY Y SY

Feature Code	Question	Answer					
	“Smells like sewer or mud”?						
C2	Does the refill drinking water "taste bitter or chewy on the tongue"?	TY	TT	KY	CY	Y	SY
C3	Is the refill drinking water “Not clear or cloudy”?	TY	TT	KY	CY	Y	SY
C4	Is the refill drinking water "A little yellowish"?	TY	TT	KY	CY	Y	SY
C5	Is the refill drinking water "Color changes in 2 days"?	TY	TT	KY	CY	Y	SY

When you want to answer a question, you will be given 6 (six) answers, each of which has the following weight:

1. TY (Not Sure) : 0
2. TT (Don't Know) : 0.2
3. KY (Not sure) : 0.4
4. CY (Sure enough) : 0.6
5. Y (sure) : 0.8
6. SY (Very Sure) : 1.0

After all the questions were answered, the MB (Measure of Belief) and MD (Measure of increased Disbelief) values were obtained for each characteristic of refill drinking water contaminated with escherichia coli bacteria as in Table 3 below:

Table 3.

Value of MB (Measure of Belief) and MD (Measure of increased Disbelief)				
No	Code	Characteristics of Water	MB value	MD value
1	C1	Smells like sewage or mud	0.8	0.2
2	C2	It feels bitter or chewy on the tongue	0.8	0.2
3	C3	Not clear or cloudy	0.8	0.2
4	C4	A little yellowish	0.6	0.4
5	C5	Color change in 2 days	0.6	0.4

Source : Medan City Environmental Health Engineering Agency

Based on the information on water characteristics and the values of MB (Measure of Belief) and MD (Measure of increased Disbelief) that have been obtained, in this case, a production rule or rule base related to Escherichia Coli bacteria is compiled, these rules are as follows:

IF smells like sewer or mud  
 AND feels bitter or chewy on the tongue  
 AND is not clear or cloudy  
 AND is a little yellowish  
 AND color changes within 2 days  
 THEN contains Escherichia coli bacteria.

After that, the calculation is done using the certainty factor method formula. And here is the formula for certainty factor methods:

$$CF(H,E) = MB(H,E) - MD(H,E) \dots \dots \dots (2.1)$$

$$CF(H,E)1 = CF(H)*CF(E).....(2.5)$$

$$CfcombineCF(H,E)1,2 = CF(H,E)1 + CF(H,E)2*(1CF(H,E)1).....(2.6)$$

$$CfcombineCF(H,E)old3 = CF(H,E)old + CF(H,E)3*(1CF(H,E)old).....(2.6)$$

Suppose the user chooses the answer to the question from the characteristics of refilled drinking water as follows:

1. Smells like sewage or mud: very sure
2. Tasting or chewing on the tongue: sure
3. Not clear or cloudy: sure
4. A little yellowish: sure enough
5. Color change in 2 days: not sure

The first step, to determine the CF value for each water characteristic is as follows:

CF (Smells like sewage or mud)

$$CF(H,E) = MB(H,E) - MD(H,E) \\ = 0.8 - 0.2 = 0.6$$

CF (It feels bitter or chewy on the tongue)

$$CF(H,E) = MB(H,E) - MD(H,E) \\ = 0.8 - 0.2 = 0.6$$

CF (Not clear or cloudy)

$$CF(H,E) = MB(H,E) - MD(H,E) \\ = 0.8 - 0.2 = 0.6$$

CF (A little yellowish)

$$CF(H,E) = MB(H,E) - MD(H,E) \\ = 0.6 - 0.4 = 0.2$$

CF (Color change in 2 days)

$$CF(H,E) = MB(H,E) - MD(H,E) \\ = 0.6 - 0.4 = 0.2$$

Furthermore, based on the user's choice of answers, the initial rule or rule which has 5 (five) premises (symptoms) is broken down into a rule that has a single premise, so that it becomes:

Rule 1.1

IF Smells like sewer or mud  
 THEN contains Escherichia coli bacteria

Rule 1.2

IF Feels tart or chit on the tongue  
 THEN contains Escherichia coli bacteria

Rule 1.3

IF Not clear or cloudy  
 THEN contains Escherichia coli bacteria

Rule 1.4

IF A little yellowish  
 THEN contains Escherichia coli bacteria

Rule 1.5

IF Color changes in 2 days  
 THEN contains Escherichia coli bacteria

Then the CF value is calculated by multiplying CF (user) by CF (expert) which is then calculated:

$$CF(H,E) = CF(Expert)*CF(User)$$

$$CF 1.1 = 0.6*1.0 = 0.6$$

$$CF\ 1.2 = 0.6 * 0.8 = 0.48$$

$$CF\ 1.3 = 0.6 * 0.8 = 0.48$$

$$CF\ 1.4 = 0.2 * 0.6 = 0.12$$

$$CF\ 1.5 = 0.2 * 0 = 0$$

The final step is to combine the CF value from each of the rules, namely combine CF 1.1 with CF 1.2 with the following formula:

CFcombine (CF1, CF2) = CF1 + CF2 \* (1 - CF1), so that becomes

$$\begin{aligned} CFcombine (CF1, CF2) &= 0.6 + 0.48 * (1 - 0.6) \\ &= 0.6 + 0.192 \\ &= 0.792\ CFold \end{aligned}$$

Combine CFold with CF 1.3

$$\begin{aligned} CFcombine (CFold, CF3) &= 0.792 + 0.48 * (1 - 0.792) \\ &= 0.792 + 0.09984 \\ &= 0.89184\ CFold2 \end{aligned}$$

Combine CFold2 with CF 1.4

$$\begin{aligned} CFcombine (CFold2, CF4) &= 0.89184 + 0.12 * (1 - 0.89184) \\ &= 0.89184 + 0.0129792 \\ &= 0.9048192\ CFold3 \end{aligned}$$

Combine CFold3 with CF 1.5

$$\begin{aligned} CFcombine (CFold3, CF5) &= 0.9048192 + 0 * (1 - 0.9437568) \\ &= 0.9048192 + 0 \\ &= 0.9048192\ CFold4 \end{aligned}$$

From the calculation above, the CF value is 0.9048192.

$$\begin{aligned} \text{Then the percentage of confidence} &= CFold4 * 100\% \\ &= 0.9048192 * 100\% \\ &= 90.48192\% \end{aligned}$$

$$\text{Can be rounded up to} \quad = 90.48\%.$$

Thus it can be said that the calculation of certainty factors in escherichia coli bacteria in refill drinking water has a percentage confidence level of 90.48% confidence.

#### 4. CONCLUSION

After conducting the discussion, the researchers took the following conclusions To diagnose bacteria in refilled drinking water, namely by determining the CF value of each of the characteristics of refilled drinking water containing Escherichia coli (e-coli) bacteria and diagnosing Escherichia coli bacteria by showing the value of the level of confidence. The application of the certainty factor method is very suitable for use in expert systems to diagnose Escherichia coli bacteria in refilled drinking water, so that the results of the level of certainty in the diagnosis of escherichia coli bacteria can be seen, which the user of this application does not know whether the refill drinking water to be consumed contains bacteria escherichia coli or not. The certainty factor method and Visual Basic.net 2008 programming can help to determine the presence of escherichia coli bacteria in refill drinking water based on the visible characteristics of refilled drinking water and as a means of preventing water containing Escherichia coli bacteria from being consumed, so as not to cause disease.

#### REFERENCES

- [1] K. E. Setyaputri, A. Fadlil, and S. Sunardi, "Analisis Metode Certainty Factor pada Sistem Pakar Diagnosa Penyakit THT," *J. Tek. Elektro*, vol. 10, no. 1, 2018, doi: 10.15294/jte.v10i1.14031.

- [2] K. Muludi, R. Suharjo, A. Syarif, and F. Ramadhani, "Implementation of forward chaining and certainty factor method on android-based expert system of tomato diseases identification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 9, 2018, doi: 10.14569/ijacsa.2018.090957.
- [3] Z. Ge, "Process Data Analytics via Probabilistic Latent Variable Models: A Tutorial Review," *Ind. Eng. Chem. Res.*, vol. 57, no. 38, 2018, doi: 10.1021/acs.iecr.8b02913.
- [4] A. S. Soma and T. Kubota, "Landslide susceptibility map using certainty factor for hazard mitigation in mountainous areas of ujung-loe watershed in South Sulawesi," *For. Soc.*, vol. 2, no. 1, 2018, doi: 10.24259/fs.v2i1.3594.
- [5] A. A. S. Nugraha, N. Hidayat, and L. Fanani, "Sistem Pakar Diagnosis Penyakit Kucing Menggunakan Metode Naive Bayes – Certainty Factor Berbasis Android," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 2, 2018.
- [6] R. Annisa, "Sistem Pakar Metode Certainty Factor Untuk Mendiagnosa Tipe Skizofrenia," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 3, no. 1, 2018.
- [7] R. Hamidi, H. Anra, and H. S. Pratiwi, "Analisis Perbandingan Sistem Pakar Dengan Metode Certainty Factor dan Metode Dempster-Shafer Pada Penyakit Kelinci," *J. Sist. dan Teknol. Inf.*, vol. 5, no. 2, 2017.
- [8] R. Oktapiani, "Penerapan Metode Forward Chaining Pada Sistem Pakar Kerusakan Komputer," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 02, no. 02, 2017.
- [9] E. R. Ritonga and M. D. Irawan, "SISTEM PAKAR DIAGNOSA PENYAKIT PARU-PARU," *J. Comput. Eng. Syst. Sci.*, vol. 2, no. 1, 2017, doi: <https://doi.org/10.24114/cess.v2i1.7179>.
- [10] N. Aini, R. Ramadiani, and H. R. Hatta, "Sistem Pakar Pendiagnosa Penyakit Tuberkulosis," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 12, no. 1, 2017, doi: 10.30872/jim.v12i1.224.