In-Silico Analysis of Chemical Compounds *Cinnamomum verum* for Antibacte

Wulandari Agustin[1], Ani Aguswita[2], Anggi Delvia Safitri[3]

[1]Universitas Negeri Padang,[2] Universitas Andalas, [3]Universitas Riau

[*]wulandaragustin@gmail.com

Abstract. Cinnamom or *Cinnamomum verum* is a natural substance that has been known as one of the spices,but later known as tradisional medicine. Cinnamon bark contains several antibacterial compounds,such as eugenol and cinnamaldehyde. Antibacteri are susbtances that can interfere with growth or even kill bacteria ny means of harmful microbial metabolism. This literature review aims to identify and analyze the trends, datasets, methods and frameworks used in the topic of attribute independence assumption assumptions on NB between 2010 and 2018. Based on the inclusion and exclusion criteria designed, it shows 71 study studies of attribute independence assumptions on the published NB between January 2010 and December 2018 are investigated in this literature review have been conducted as a review of systematic literature. A systematic literature review is defined as the process of identifying, assessing, and interpreting all available research evidence in order to provide answers to specific research questions.

1. Introduction

Cinnamon *(Cinnamomum verum)* is a type of spices commonly used for cooking,but cinnamon is also traditionally used as a supplement for various diseases,with honey,for example for the treatment of arthritis,skin,heart,and flatulence. And other benefits of cinnamon are for antibacterial,this is because the dominant chemical component contained in essesntial oils cinnamomum verum is cinamaldehyde,eugenol and camphor. Sinamaldehyde and eugenol is the composition that has the greates pharmacologies effect.

Cinnamon has various beefits as a pharmacological effect,wrong one is as antibacterial. The mechanism of action of sinamaldehyde is able inhibits the synthesis of β-(1,3)-glucan and chitin which are the main components of the bacteri cell wall. Meanwhile one of the works eugenol is inhibiting the biosynthesis of ergosterol which is a major constituent trichophtyton rubrum cell membrane. Apart from having an antibacterial effect, Cinnamon also has other benefits,namely anti-oxidants,anti-inflammatory,anti-diabetes,anti-fungal,insecticide and analgesic.

The fruit,steams,leaves and rooth of cinnamon have different chemical compositions. The difference can be seen in the Table 1.

Table 1. Chemical components of the plant parts of Cinnamom Verum

Leaves	• Cinnamaldehyde • Eugenol
Bark	• Cinnamaldehyde • Eugenol
Root bark	Camphor
Fruit	• Trans-Cinnamyl acetate • caryophyllene

2. Methodology

This paper will use the SLR approach to review research on the Naïve Bayes algorithm with the problem of attribute independence assumptions. Systematic Literature Review (SLR) is a process for identifying, assessing, and interpreting all available research with a view to providing answers to specific RQs[15]. In the guide that Kitchenham has made in 2007[15], the literature review will be compiled based on the Systematic Literature Review. Conclusions described. The last step, synthesize the data in which the purpose of collecting evidence from the survey paper that has been obtained to answer RQ.Synthesis data used in this study, will generally be a narrative synthesis. Some tables and visual tools will be used to support the explanation in this study.

Table 2. Inclusion and Exclusion Criteria

Inclusion Criteria	The study discusses the feature independence assumption on the naïve bayes algorithm
	For research that has two types of journal and conference publications, it will take the type of journal publication
	For duplicate research, the most comprehensive and up-to-date data is available
Exclusion Criteria	The study discusses the independent attribute assumption on naïve nayes method, but does not propose methods to overcome the independent attribute assumption on the naïve bayes method
	Research that does not use strong validation
	Studies not written in English

Table 3. Data Extraction for RQ

Data Extraction	Research Question (RQ)
Research and Year of Publication	RQ1, RQ2
A frequently used dataset	RQ3
Method approach is often used	RQ4

3. Result and Discussion

As mentioned earlier, this literature review will be limited to journals published in 2010 through 2018. The time span is to see if research on the feature independence assumption on the Naïve Bayes method is still relevant. In Figure 3 it can be seen that the trend of research from 2010 to 2016 has increased, so it can be concluded that research on the assumption of attribute independence on the Naïve Bayes method is still very relevant to date.'

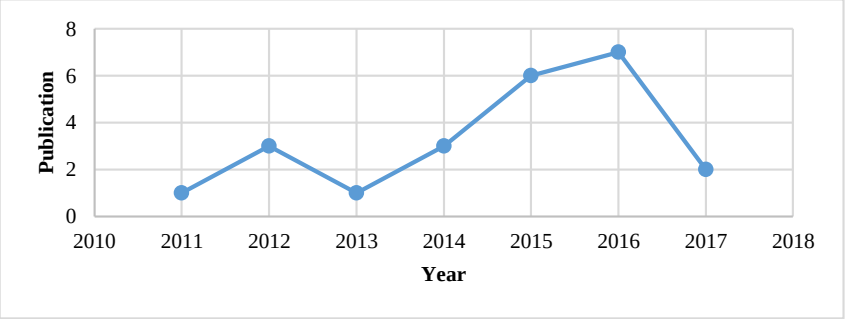

Figure 1. Publication Per-Year

Then in Figure 4 can be seen a journal that published a paper about the assumption of feature independence in the Naïve Bayes method. For the record, the journal in question is a journal that publishes papers that have been selected.

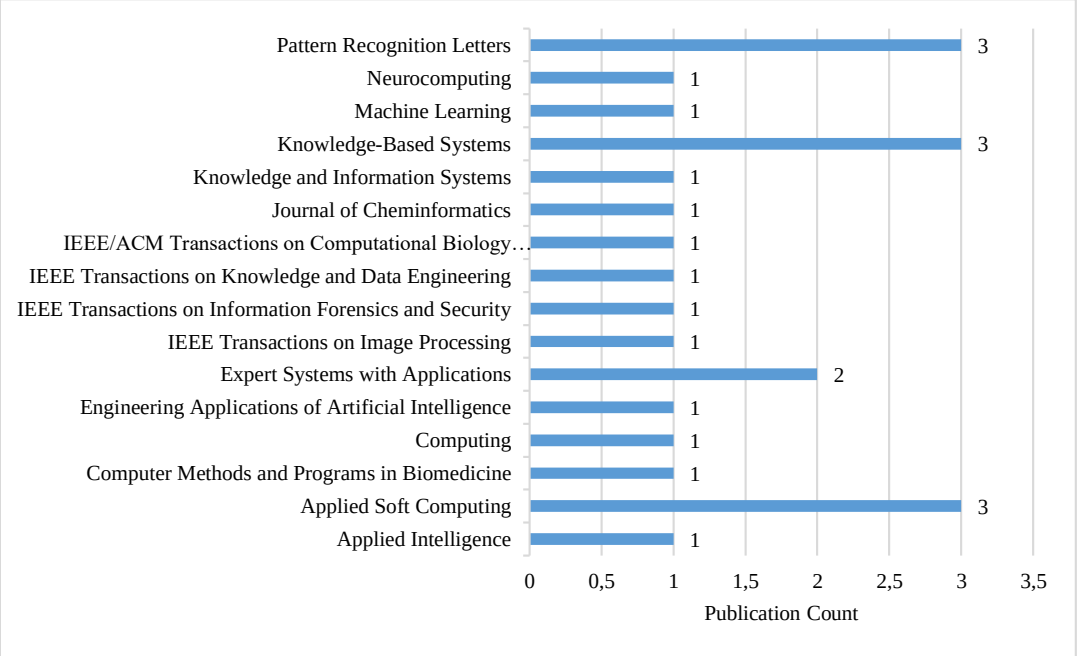

Figure 2. Number of publications per year

Contributing to the research topic on the assumption of attribute independence on the Naïve Bayes method will be investigated and identified. Figure 5 shows the most active researcher on research on the assumption of attribute independence on the Naïve Bayes method. The most influential researchers were Liang Xiao Jiang and followed by Lungan Zhang, and Shasa Wang. In addition it is not the first writer that is Chaoqun Li and Yang Xiang.

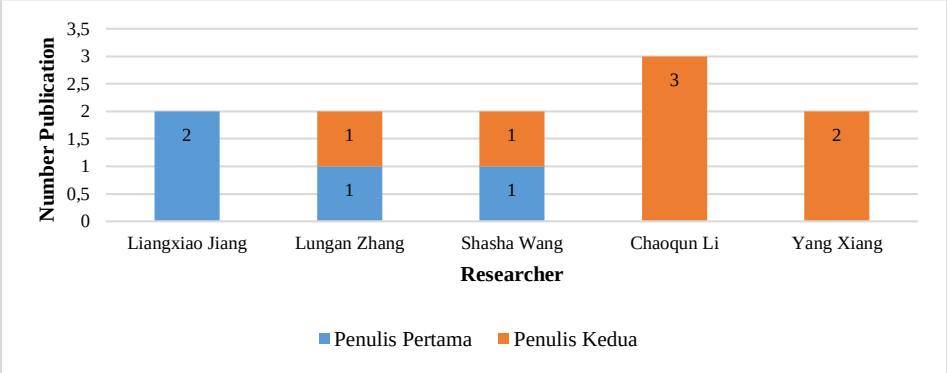

Figure 3. The Most Active and Influential Researcher

Figure 6 shows the percentage of total data sets used from 2010 to 2018. 70% of researchers use public datasets and 30% of researchers use private datasets. Public datasets are mostly accessible at the University of California Irvine (UCI) repository[16].

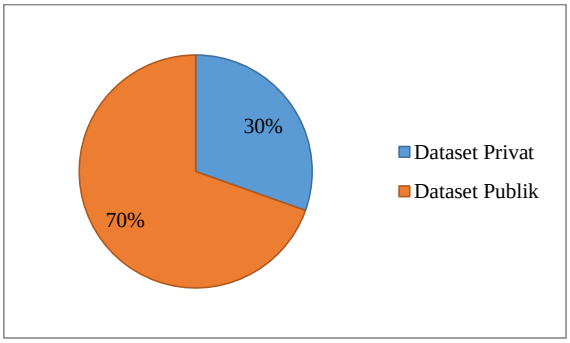

Figure 4. Dataset Distribution

The main study distribution over the years, and by source, is presented in Figure 7. More research has been published, and more public datasets have been used for research topics assuming attribute independence in naïve bayes since 2011.

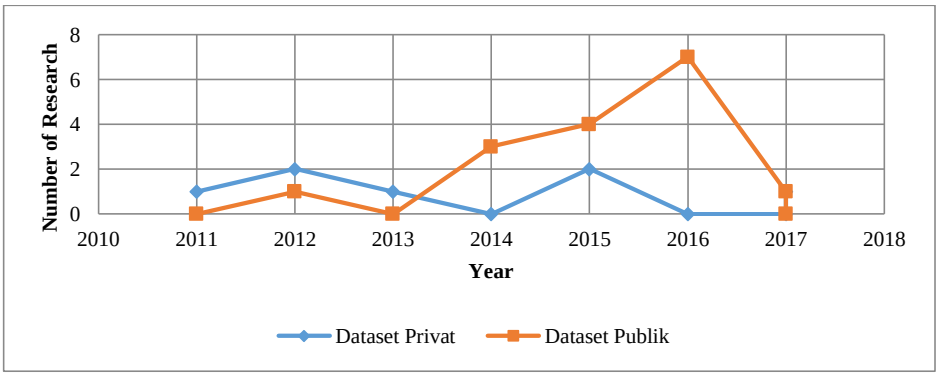

Figure 5. Distribution of private datasets and public datasets

On investigation, there are three strategies method approaches used to overcome the assumption of independence in naive bayes, including: 1) weighted strategies based on single correlation (Mutual Information (MI) [2], Attribute Weighted K-Nearest Neighboard (AWKNN) [17], Hidden Naïve Bayes (HNB) [3], Attribute weighted Naive Bayes using mutual information weighted method (MIWNB) [4], GRWNB [18]), 2) attribute weighting strategy using attribute correlation (like

CFSWNB [19], SBC [20], TreeWNB [21], ReFWNB [22], FDNB [23]), and 3) self-adaptive attribute strategy (like NACO [24], DPGA [25], ES [26], SODE [27] and AISWNB [28]).

Some researchers propose several techniques to improve the accuracy of previously proposed classifier to overcome attribute independence on NB. This proposed technique has recently attempted to improve the prediction accuracy of methods generated by the modification and incorporation of several machine learning methods [27], add feature selection method [17] using several methods of optimizing evolutionary calculations [29].

Sixteen different methods have been applied to find the attribute independence solution on the NB method. Of the sixteen methods are found the most frequently used method of Mutual Information (MI) [2], metode Selective Bayes Classifier (SBC) [20], and Immune Systems based weighting scheme for Naive Bayes classification (AISWNB) method [28].

4. **Conclusion**

This literature review aims to identify and analyze the trends, datasets, methods and frameworks used in the topic of attribute independence assumption assumptions on NB between 2010 and 2018. Based on the inclusion and exclusion criteria designed, it shows 71 study studies of attribute independence assumptions on the published NB between January 2010 and December 2018 are investigated in this literature review have been conducted as a review of systematic literature. A systematic literature review is defined as the process of identifying, assessing, and interpreting all available research evidence in order to provide answers to specific research questions.

The results of this study identified three of the most commonly used and influential framework methods in the topic of attribute independence on the NB. They are Menzies et al. Framework, Lessmann et al. Framework, and Song et al. Framework. They are Langley et al [20], Friedman et al [2], and Wu et al [28].

References

[1] J. Hernández-González, I. Inza, and J. A. Lozano, "Learning Bayesian network classifiers from label proportions," *Pattern Recognit.*, vol. 46, no. 12, pp. 3425–3440, 2013.

[2] N. Friedman, D. Geiger, M. Goldszmidt, G. Provan, P. Langley, and P. Smyth, "Bayesian Network Classifiers *," *Mach. Learn.*, vol. 29, pp. 131–163, 1997.

[3] L. Jiang, H. Zhang, and Z. Cai, "A novel bayes model: Hidden naive bayes," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 10, pp. 1361–1371, 2009.

[4] L. Jiang, H. Zhang, Z. Cai, and D. Wang, "Weighted average of one-dependence estimators," *J. Exp. Theor. Artif. Intell.*, vol. 24, no. 2, pp. 219–230, 2012.

[5] P. harliana and R. Rahim, "Comparative Analysis of Membership Function on Mamdani Fuzzy Inference System for Decision Making," *J. Phys. Conf. Ser.*, vol. 930, no. 1, p. 012029, Dec. 2017.

[6] R. Rahim *et al.*, "Searching Process with Raita Algorithm and its Application," *J. Phys. Conf. Ser.*, vol. 1007, no. 1, p. 012004, Apr. 2018.

[7] R. Rahim, A. S. Ahmar, A. P. Ardyanti, and D. Nofriansyah, "Visual Approach of Searching Process using Boyer-Moore Algorithm," *J. Phys. Conf. Ser.*, vol. 930, no. 1, p. 012001, Dec. 2017.

[8] R. Rahim, S. Nurarif, M. Ramadhan, S. Aisyah, and W. Purba, "Comparison Searching Process of Linear, Binary and Interpolation Algorithm," *J. Phys. Conf. Ser.*, vol. 930, no. 1, p. 012007, Dec. 2017.

[9] R. Rahim, Nurjamiyah, and A. R. Dewi, "Data Collision Prevention with Overflow Hashing Technique in Closed Hash Searching Process," *J. Phys. Conf. Ser.*, vol. 930, no. 1, p. 012012, Dec. 2017.

[10] R. Rahim, D. Hartama, H. Nurdiyanto, A. S. Ahmar, D. Abdullah, and D. Napitupulu, "Keylogger Application to Monitoring Users Activity with Exact String Matching Algorithm,"

J. Phys. Conf. Ser., vol. 954, no. 1, p. 012008, 2018.
[11] C. Zhang, G.-R. Xue, Y. Yu, and H. Zha, "Web-scale classification with naive bayes," *Proc. 18th Int. Conf. World wide web - WWW '09*, p. 1083, 2009.
[12] R. S. Wahono, "A Systematic Literature Review of Software Defect Prediction : Research Trends , Datasets , Methods and Frameworks," *J. Softw. Eng.*, vol. 1, no. 1, pp. 1–16, 2015.
[13] C. Catal and B. Diri, "A systematic review of software fault prediction studies," *Expert Syst. Appl.*, vol. 36, no. 4, pp. 7346–7354, 2009.
[14] A. S. Ahmar *et al.*, "Modeling Data Containing Outliers using ARIMA Additive Outlier (ARIMA-AO)," *J. Phys. Conf. Ser.*, vol. 954, no. 1, 2018.
[15] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3," *Engineering*, vol. 45, no. 4ve, p. 1051, 2007.
[16] D. Aha *et al.*, "UCI Repository of Machine Learning Database," 1987.
[17] J. Wu, Z. Cai, S. Zeng, and X. Zhu, "Artificial immune system for attribute weighted Naive Bayes classification," *Proc. Int. Jt. Conf. Neural Networks*, no. 61075063, 2013.
[18] H. Zhang and S. Sheng, "Learning weighted naive bayes with accurate ranking," *Proc. - Fourth IEEE Int. Conf. Data Mining, ICDM 2004*, pp. 567–570, 2004.
[19] M. A. Hall, "uow-cs-wp-2000-08.pdf." 2000.
[20] P. Langley and S. Sage, "Induction of Selective Bayesian Classifiers," *Proc. Tenth Int. Conf. Uncertain. Artif. Intell.*, no. 1990, pp. 399–406, 1994.
[21] M. A. Hall, "A decision tree-based attribute weighting filter for naive Bayes.pdf," 2006.
[22] M. Robnik-Siknja and I. Kononeko, "Theoretical and empirical analysis of RelifF and RReliefF," *Mach Learn*, vol. 53, pp. 23–69, 2003.
[23] Ö. F. Arar and K. Ayan, "A Feature Dependent Naive Bayes Approach and Its Application to the Software Defect Prediction Problem," *Appl. Soft Comput. J.*, 2017.
[24] M. Borrotti, G. Minervini, D. De Lucrezia, and I. Poli, "Naïve Bayes ant colony optimization for designing high dimensional experiments," *Appl. Soft Comput. J.*, pp. 1–10, 2016.
[25] T. Park and K. R. Ryu, "A dual-population genetic algorithm for adaptive diversity control," *IEEE Trans. Evol. Comput.*, vol. 14, no. 6, pp. 865–884, 2010.
[26] Y. P. Huang, Y. T. Chang, S. L. Hsieh, and F. E. Sandnes, "An adaptive knowledge evolution strategy for finding near-optimal solutions of specific problems," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3806–3818, 2011.
[27] J. Wu, S. Pan, X. Zhu, P. Zhang, and C. Zhang, "SODE: Self-Adaptive One-Dependence Estimators for classification," *Pattern Recognit.*, vol. 51, pp. 358–377, 2016.
[28] J. Wu, S. Pan, X. Zhu, Z. Cai, P. Zhang, and C. Zhang, "Self-adaptive attribute weighting for Naive Bayes classification," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1487–1502, 2015.
[29] J. Lakoumentas, J. Drakos, M. Karakantza, G. Sakellaropoulos, V. Megalooikonomou, and G. Nikiforidis, "Optimizations of the naïve-Bayes classifier for the prognosis of B-Chronic Lymphocytic Leukemia incorporating flow cytometry data," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 158–167, 2012.