Trend of Supervised LearningModels Based Articles

Lukas Priyambodo, Padang College, Jakarta.

Abstract:

Supervised learning techniques also known as classification techniques are an important aspect of data mining techniques that assist us in classifying the object based upon certain additional information like labels. The objective of this study is to present the Google Scholar's publication trend of supervised learning based article published in the domain of marketing, healthcare, text processing as well as in agriculture. The statistics reveal that most of the supervised learning techniques based article has been published in healthcare followed by marketing.

Keywords: supervised learning techniques; marketing, healthcare, text processing, agriculture and data mining.

1. Introduction

Supervised learning technique is also known as classification techniques are important part of data mining techniques. These techniques are normally used to classify objects into binary or multiple classes. The case where the data is classified into two classes is known as binary classification and where more than two classes exist, such classification process is known as multi-classification.

In last few years, lot of research work has been done using supervised learning techniques in the area of agriculture[1][2][3], marketing [4][5], healthcare [6][7][8][9][10][11][12][13] and text processing [14][15]. Several authors have used various classifiers like naïve Bayes, decision tree, support vector machine, rule-based classifiers and various stochastic algorithms like genetic algorithm, ant colony optimization, artificial bee colony, firefly algorithm, neural network and deep learning techniques[16][17].

Data mining has been instigated from three different techniques viz. Statistics, Artificial Intelligence and Machine Learning. Several heuristics have been projected to perk up the competence of the data mining process. , clustering, association mining and prediction are four major tasks of data mining technique. In general, classification is categorized as single or multi-class. In single class, there is only one class label that has to be recognized. The elements that belong to the class are known as normal and rest of the elements are categorized as anomalies [6].

2. Literature Review

I have studied the role of supervised learning models in the area of agriculture, marketing, healthcare and text processing. It has been found that different researchers have used different supervised learning techniques for mining data in these areas. In agriculture, people work on finding the relationship between spray and food/vegetables, Prediction of problematic wine fermentation, plants disease diagnosis, optimizing pesticides etc [18][19][20]. Likewise, in healthcare, different human disease has been diagnosed [21][22][23]. Moreover, some of the people have also find the association between medicine

and health of the person. In-text processing, opinion mining, web mining and sentiment analysis are on the top list[24][25].

3. Publication Trend

The major intention of this research work is to highlight the publication trend of supervised learning techniques based articles published in the four major areas i.e. agriculture, healthcare, marketing as well as text processing. Figure 1 represents the number of supervised learning techniques based article published in this four domain. As per google scholar, the number of articles published in agriculture, healthcare, marketing as well as text processing is 605, 1460, 1450 and 623 respectively. Out of these four domains healthcare and marketing are most explored areas.
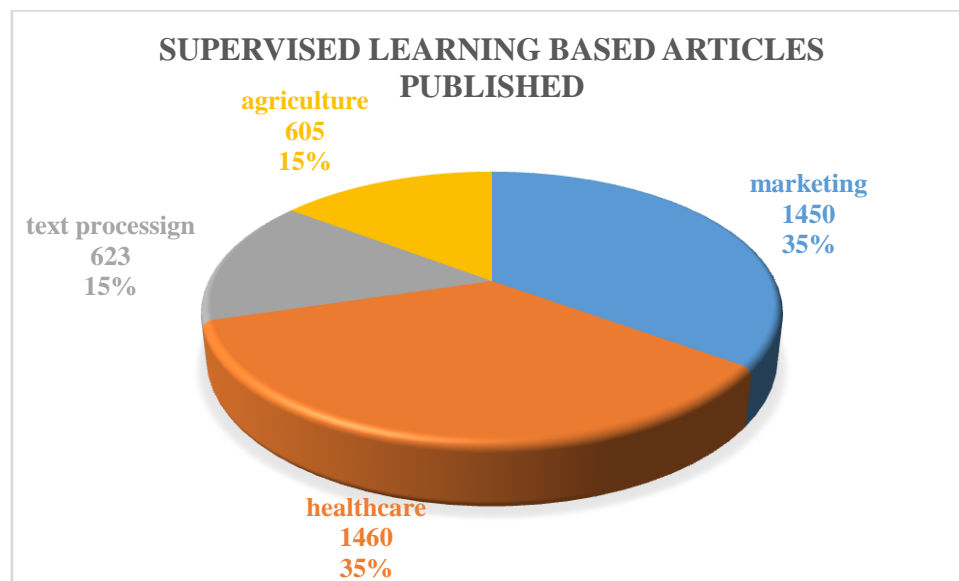

Figure 1: Supervised learning trend

In addition, the statistics of publication of supervised learning technique based articles of the last ten years (2010-2019) has also been presented in Table 1. From the data values of Table 1, it is observed that consistent growth in number of indexed articles has been increased over last ten years.

Table 1: Stats of publication of supervised learning technique based articles (2010-2019)

	Marketing	Healthcare	Text processing	Agriculture
2010	41	38	26	16
2011	48	47	32	21
2012	62	60	34	32
2013	77	72	46	32
2014	113	91	39	31
2015	245	118	61	43
2016	139	148	70	53
2017	159	215	88	69
2018	221	261	48	82
2019	192	321	57	123

Figure 2 presents the statistics of publication (2010-2019) of supervised learning based publication published in the areas of marketing, healthcare, text processing and agriculture.
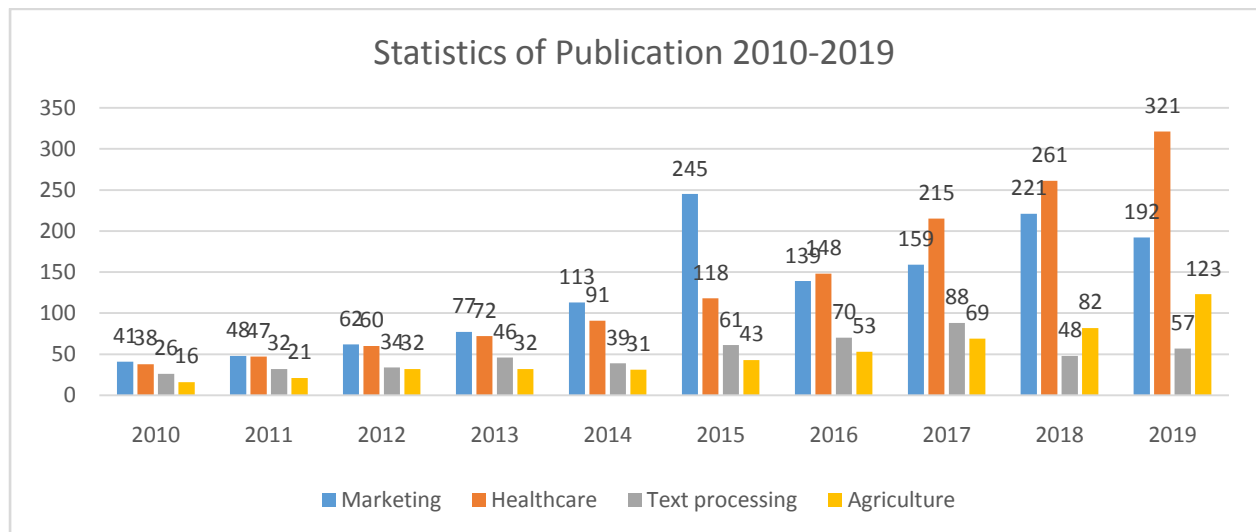

Figure 2: Statistics of Publication

4. Conclusion

The objective of this research work is to present the publication trend of supervised learning based article published in the areas of marketing, healthcare, text processing and agriculture. The last ten years of analysis of the publication trend has been examined and analyzed. It has been found that most of the work has been carried out for marketing and healthcare domain. The rate of publication of supervised learning based articles for these domains is almost double than the text processing and agriculture. The results shows that there is more scope in using supervised learning in text processing and agriculture as compared to the marketing and healthcare.

5. References

1. Corrales, David Camilo, Juan Carlos Corrales, and Apolinar Figueroa-Casas. "Towards detecting crop diseases and pest by supervised learning." Ingeniería y Universidad 19.1 (2015): 207-228.
2. Shakoor, MdTahmid, et al. "Agricultural production output prediction using supervised machine learning techniques." 2017 1st International Conference on Next Generation Computing Applications (NextComp). IEEE, 2017.
3. Liakos, Konstantinos G., et al. "Machine learning in agriculture: A review." Sensors 18.8 (2018): 2674.
4. Jeyapriya, A., and CS KanimozhiSelvi. "Extracting aspects and mining opinions in product reviews using a supervised learning algorithm." 2015 2nd International Conference on Electronics and Communication Systems (ICECS). IEEE, 2015.
5. Elsalamony, Hany A. "Bank direct marketing analysis of data mining techniques." International Journal of Computer Applications 85.7 (2014): 12-22.
6. Sharma, M., G. Singh, and R. Singh. "Stark assessment of lifestyle based human disorders using data mining based learning techniques." IRBM 38.6 (2017): 305-324.

7. Vijiyarani, S., and S. Sudha. "Disease prediction in data mining technique–a survey." International Journal of Computer Applications & Information Technology 2.1 (2013): 17-21.

8. Kaur, Prableen, and Manik Sharma. "A survey on using nature inspired computing for fatal disease diagnosis." *International Journal of Information System Modeling and Design (IJISMD)*8.2 (2017): 70-91.

9. Meng, Gilliar, and HebaSaddeh. "Performance Analysis of Different Classifier for Diabetes Diagnosis." *International Journal of Computer Applications & Information Technology*11.2 (2019): 265-270.

10. Sharma, Manik, Gurvinder Singh, and Rajinder Singh. "An Advanced Conceptual Diagnostic Healthcare Framework for Diabetes and Cardiovascular Disorders." *arXiv preprint arXiv:1901.10530* (2019).

11. Gautam, Ritu, Prableen Kaur, and Manik Sharma. "A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings." *Progress in Artificial Intelligence* (2019): 1-24.

12. Fatima, Meherwar, and Maruf Pasha. "Survey of machine learning algorithms for disease diagnostic." *Journal of Intelligent Learning Systems and Applications* 9.01 (2017): 1.

13. Diwani, Salim Amour, and Anael Sam. "Diabetes Forecasting Using Supervised Learning Techniques." *Advances in Computer Science: an International Journal* 3.5 (2014): 10-18.

14. Sharma, Manik, Gurvinder Singh, and Rajinder Singh. "Design of GA and Ontology based NLP Frameworks for Online Opinion Mining." *Recent Patents on Engineering* 13.2 (2019): 159-165.

15. Verma, Rohit, and Dr Jahid Ali. "A-survey of feature extraction and classification techniques in OCR systems." *International Journal of Computer Applications & Information Technology*1.3 (2012): 1-3.

16. Kaur, Prableen, and Manik Sharma. "Diagnosis of Human Psychological Disorders using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis." Journal of medical systems 43.7 (2019): 204.

17. Sharma, Manik, Gurvinder Singh, and Rajinder Singh. "Accurate Prediction of Life Style Based Disorders by Smart Healthcare Using Machine Learning and Prescriptive Big Data Analytics." Data-Intensive Computing Applications for Big Data29 (2018): 428.

18. Rumpf, T., et al. "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance." Computers and electronics in agriculture 74.1 (2010): 91-99.

19. Mahlein, Anne-Katrin, et al. "Recent advances in sensing plant diseases for precision crop protection." European Journal of Plant Pathology 133.1 (2012): 197-209.

20. Cortez, Paulo, et al. "Modeling wine preferences by data mining from physicochemical properties." Decision Support Systems 47.4 (2009): 547-553.

21. Ghorbani, R., and R. Ghousi. "Predictive data mining approaches in medical diagnosis: A review of some diseases prediction." International Journal of Data and Network Science3.2 (2019): 47-70.

22. Arshad, Insha, et al. "Liver disease detection due to excessive alcoholism using data mining techniques." 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE). IEEE, 2018.

23. Singh, Poornima, Sanjay Singh, and Gayatri S. Pandi-Jain. "Effective heart disease prediction system using data mining techniques." International journal of nanomedicine 13.T-NANO 2014 Abstracts (2018): 121.

24. Miner, Gary, et al. Practical text mining and statistical analysis for non-structured text data applications. Academic Press, 2012.

25. Dey, Lipika, and SK MirajulHaque. "Opinion mining from noisy text data." International Journal on Document Analysis and Recognition (IJDAR) 12.3 (2009): 205-226.