# A Case Study of Medical Big Data Processing: Data Mining for the Hyperuricemia

# Abstract


- Today, the world comes into a booming information age. With the rapid development of computer and sensors, various bits of data are continuously being generated as time goes by. The amount of data to be processed has entered into the big data category.
- Data mining is widely used to discover hidden information in the large amounts of data.
- However, data mining applied to medical databases is a challenging process. The unavailability of large raw data and data complexity are some of the difficulties encountered.

# Contd..

- This research work proposes a way of dealing with big medical data with a case study of gout disease. Gout is common chronic disease caused by the most important risk factor hyperuricemia.
- There is no drug to completely cure the gout, the patient is suffering a lot of pain. It is important to control the occurrence of gout and to study the association of gout with other metabolic diseases.
- This paper discusses methods for the analysis of this complex dataset of this disease, to help get more understanding of the disease and associated diseases.

# Existing

- In today's booming information age, with the rapid development of computer and sensors, various bits of data are continuously being generated as time goes by The amount of data to be processed is more and more large, and the number of data has entered into the big data category.
- However, how to effectively organize and use these data has become a problem.
- Aside from what it says on the surface, data also contain hidden information that may bring major business opportunities.
- For various industries, it's necessary to discover the hidden information through processing these data, so data mining has been widely studied in many industries.

# Cond…

- The amount of data to be processed has entered into the big data category.
- Data mining is widely used to discover hidden information in the large amounts of data.
- However, data mining applied to medical databases is a challenging process. The unavailability of large raw data and data complexity are some of the difficulties encountered.
- This research work proposes a way of dealing with big medical data with a case study of gout disease. Gout is common chronic disease caused by the most important risk factor hyperuricemia.
- There is no drug to completely cure the gout, the patient is suffering a lot of pain. It is important to control the occurrence of gout and to study the association of gout with other metabolic diseases

# Disadvantage

- However, data mining applied to medical databases is a challenging process. The unavailability of large raw data and data complexity are some of the difficulties encountered.
- Gout is common chronic disease caused by the most important risk factor hyperuricemia. There is no drug to completely cure the gout, the patient is suffering a lot of pain.
- It is important to control the occurrence of gout and to study the association of gout with other metabolic diseases

# Proposed

- In this paper, data mining is studied for the application of medical data.
- Based on the clinical data of hyperuricemia, all medical data has been preprocessed, classified, and mined in order to find the relationship between hyperuricemia disease and other common chronic diseases and propose an associated relation which applies auxiliary support for doctor's clinical diagnosis and disease research.

# Advantages

- By data mining, people can find information, regularity and producing value from the data.
- An associated relation is proposed according to the experiments, which applies auxiliary support for doctor's clinical diagnosis and disease research in the local.

# HARDWARE REQUIREMENTS

- Processor – Pentium –III
- Speed – 1.1 Ghz
- RAM – 256 MB(min)
- Hard Disk – 20 GB
- Floppy Drive – 1.44 MB
- Key Board – Standard Windows Keyboard
- Mouse – Two or Three Button Mouse
- Monitor – SVGA

# SOFTWARE REQUIREMENTS

- Operating System : Windows 8
- Front End : Java /DOTNET
- Database : Mysql/HEIDISQL

# Conclusion

- This paper mainly studies the application of mining algorithm in medical large data. The relationship between common metabolic diseases and hyperuricemia was studied by using Apriori Association Rules and Logistic regression analysis based on the data of hyperuricemia.
- By comparing two algorithms, it can be seen that hypertension, Cr, TG, BUN and total bilirubin are the risk factors of hyperuricemia, in which the correlation between Cr, total bilirubin and hyperuricemia is mainly due to the age and regional limitations of data sources.
- From the results of the experiment, it can be seen that the correlation between hyperuricemia and other chronic metabolic diseases, in addition to the determinants of medical relevance, depends on the source of the data.

# Contd ..

- the age and region of the tester of the data source. Therefore, the excavation of different region and age of medical data has a great help to doctors in the treatment of disease.
- in the next stage of the study, we will focus on data such as different age, gender, and environmental factors to establish a more perfect model of hyperuricemia and other related disease through data mining

# Reference


[1] F. Padillo, J. M. Luna and S. Ventura, "An evolutionary algorithm for mining rare association rules: A Big Data approach," 2017 IEEE Congress on Evolutionary Computation (CEC), Donostia, San Sebastián, Spain, 2017, pp. 2007-2014.

[2] J. HAN, M. KAMBER, "Data Mining Concepts and Techniques", California:Morgankaufmann Publishers, 2011, pp.149-167.

[3] L. Liu, Y. Xu, Z. Ma. "A Survey to data Preprocessing in data Mining", Journal of Gansu Sciences, 2003

[4] Koh, Hian Chye, and Gerald Tan. "Data mining applications in healthcare", Journal of Healthcare Information ManagementVol 19.2 (2011):65.

[5] J. L. Breault, C. R. Goodall and P. J. Fos, "Data mining a diabetic data warehouse", Artificial Intelligence in Medicine,2002, 26(1), 37-54.