ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

A Survey on Privacy Preserving Patient-Centric Clinical Decision Support System Using Naive Bayesian Classifier

Pradnya V Kul, Dr. V. S. Bidve

M.E Student, Dept, of I.T, SKN College of Engineering, Pune, India

Associated Professor, Dept. of I.T., SKN College of Engineering, Pune, India

ABSTRACT:CDS is the term used to describe information presented at the appropriate time to enable providers and their patients to make the best decision based on the specific circumstances. By comparing the information in a patient's electronic record with a set of evidence-based clinical guidelines, an electronic CDS system.The advantages of clinical decision support system include not only improving diagnosis accuracy but also reducing diagnosis time. Specifically, with large amounts of clinical data generated every day, naive Bayesian classification can be utilized to excavate valuable information to improve clinical decision support system.It is privacy-preserving patient-centric clinical decision support system, which helps clinician complementary to diagnose the risk of patients' disease in a privacy-preserving way. The past patients' historical data are stored in cloud and can be used to train the na¨ıve Bayesian classifier without leaking any individual patient medical data, and then the trained classifier can be applied to compute the disease risk for new coming patients and also allow these patients to retrieve the top-k disease names according to their own preferences.

KEYWORD: Clinical Decision Support System; Privacy-Preserving,patient-centric, clinical decision support system.

I. INTRODUCATION

Clinical Decision System (CDS) is the term used to describe information presented at the appropriate time to enable providers and their patients to make the best decision based on the specific circumstances. By comparing the information in a patient's electronic record. To speed up the diagnosis time and improve the diagnosis accuracy, a new system in healthcare industry should be workable to provide a much cheaper and faster way for diagnosis. Naïve Bayesian Classifier, one of the popular machine learning tools, has been widely used recently to predict various diseases in Clinical Decision Support System (CDSS). It perform well in multi class prediction The past patients' historical data are stored in cloud and can be used to train the naive Bayesian classifier without leaking any individual patient medical data and then the trained classifier can be applied to compute the disease risk for new coming patients and also, allow these patients to retrieve the top-k disease names according to their own preferences.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

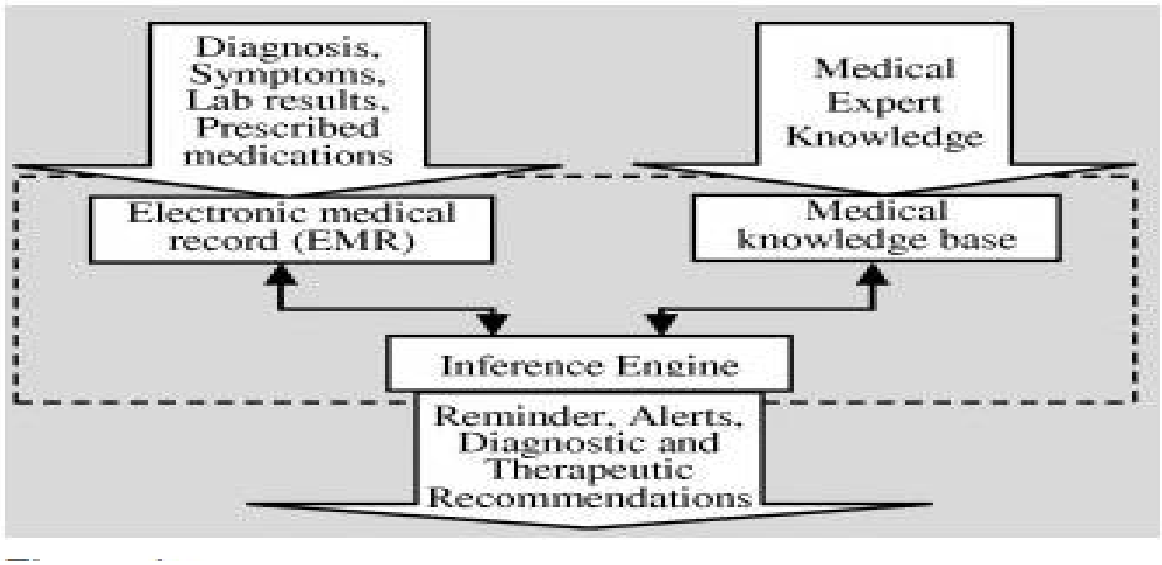

Figure I: CDSS Architecture

CDSS having two types of data first is historical data and second is current information. On the bases of that two information Naïve Bayesian classification is working, it gives necessary treatment for patient.

II. LITERATURE SURVEY

Yogachandran Rahulamthavan, Suresh Veluru, Raphael C, Jonathon A, Muttukrishannan Rajarajan L: These author has investigate "Privacy-Preserving Clinical Decision Support System using Gaussian Kernel Based Classification". It describes the cloud computing technology which is having very rich clinical data. Using this system it enhances the decision making ability of health care professionals. For this purpose it uses paillier cryptosystem but it only encrypt integer value only depends on locally available clinical data.Gaussian kernel only works purely in plain domain and it can't be modified to clinician server

ErmanAyday, Jean Louis Raisaro, Paul J. McLaren, Jacques Fellay, Jean-pierreHubaux:These author has investigate "Privacy-Preserving Computation of Disease Risk by Using Genomic, Clinical, and Environmental Data". It describes privacy for storing and processing unit in system. For this it uses homomorphic encryption and Privacy-Preserving integer comparison. It uses real patient's data and reliable disease risk factor.It works efficiently only for genomic data. It specify disease risk test using genomic data.

H. Monkaresi, R. A. Calvo, and H. Yan: These authors introduced "A Machine Learning Approach to ImproveContactless Heart Rate Monitoring Using a Webcam". We have evaluated a method for remote HR measuring in
three applications: a controlled laboratory task, a naturalisticHCI, and an indoor cycling exercise. This study evaluated Poh*et al.*'s method and showed the feasibility of their methodologyto measure HR at rest.

Rongxing Lu, Hui Zhu, Ximeng Liu, Joseph K. Liu, and Jun Shao: These authors studied "Toward Efficient and Privacy-Preserving Computing in Big Data Era". This paper has investigated the privacy challenges in thebig data era by first identifying big data privacy requirements andthen discussing whether existing privacy-preserving techniquesare

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

sufficient for big data processing. They have also introducedan efficient and privacy-preserving cosine similarity computingprotocol in response to the efficiency and privacy requirementsof data mining in the big data era.

Y. Tong, J. Sun, S. S. M. Chow, and P. Li: This paper introduced "Cloud-Assisted Mobile-Access of Health Data withPrivacy and Auditability". The author has proposed to build privacy into mobile healthsystems with the help of the private cloud. We provided a solutionfor privacy-preserving data storage by integrating a PRFbased
key management for unlikability, a search and accesspattern hiding scheme based on redundancy, and a secure indexingmethod for privacy-preserving keyword search. We alsoinvestigated techniques that provide access control (in both normaland emergency cases) and auditability of the authorizedparties to prevent misbehaviour, by combining ABE-controlledthreshold signing with role-based encryption.

Tien Tuan AnhDinh, AnwitamanDatta: This paper has introduced "Stream on the Sky: Outsourcing Access Control Enforcement for Stream Data to the Cloud". In this paper, has presented a system providing fine-grained access control for stream data over untrusted clouds. This system allows the owners to encrypt data before relaying them to the cloud. Encryption ensures both confidentiality against the cloud and access control against dishonest users. Stream force uses combinations of three encryption schemes: a deterministic scheme, a proxy ABE scheme and a sliding-window scheme. We have showed how the cloud can enforce access control over ciphertexts by transforming them for authorized user, without learning the plaintexts.

Y. Elmehdwi, B. K. Samanthula, and W. Jiang: Authors have introduced "Secure k-Nearest Neighbor Query over Encrypted Data in Outsourced Environment". This paper has proposed two novelSkNN protocols over encrypted data in the cloud. The firstprotocol, which acts as a basic solution, leaks some informationto the cloud. On the other hand, our second protocol isfully secure, that is, it protects the confidentiality of the data,user's input query, and also hides the data access patterns.

Y. Rahulamathavan, S. Veluru, R. Phan, J. Chambers, and M. Rajarajan: Authors have introduced "Privacy-Preserving Clinical Decision Support System using Gaussian Kernel based Classification". This paper has proposed a privacy-preserving decision support system using a Gaussian kernel based support vector machine. Since the proposed algorithm is a potential application of emerging outsourcing techniques such as cloud computing technology, rich clinical data sets (or healthcare knowledge) available in remote locations could be used by any clinicians via the Internet without compromising privacy, thereby enhancing the decision making ability of healthcare professionals.

J. Chen, H. Huang, S. Tian, and Y. Qu: The name of this paper is "Feature selection for text classification with Naïve Bayes". This paper presents two feature evaluation metrics (CDM and MOR) for the Naïve Bayes classifier applied on multi-class text collections. They have compared CDM and MOR with EOR, WOR and MC-OR, three variations of Odds Ratio for multi-class datasets.

R. Bellazzi and B. Zupan: Theses authors have introduced "Predictive data mining in clinical medicine: Current issues and guidelines". The goal of this review is to discuss the extent and role of the research area of predictive
data mining and to propose a framework to cope with the problems of constructing,assessing and exploiting data mining models in clinical medicine.

B. K. Samanthula, Y. Elmehdwi, G. Howser, and S. Madria: These authors have introduced "A secure data sharing and query processing framework via federation of cloud computing". This paper proposes an efficient and Secure Data Sharing (SDS) framework that prevents information leakage from a pre- viously revoked user rejoining the system. The framework utilizes the Elgamal-like encryption scheme that is additive homomorphic and which exhibits proxy re-encryption properties. Also, the proposed framework is secure as per the security definition of Secure Multi-Party Computation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

X. Yi and Y. Zhang: The authors have introduced "Privacy-preserving naive Bayes classification on distributed data via semi-trusted mixers". The proposed multi-party protocol is built on the semi-trusted mixer model, in which each data site sends messages to two semi-trusted mixers, respectively, which run our two- party protocol then broadcast the classification result. Because our protocol assumes no collusion between the two mixers and does not require the communicationchannels among data sites, it facilitates both trust management and implementation.

IV.CONCLUSION

It have proposed a privacy-preserving patient-centric clinical decision support system using naive Bayesian classifier by taking the advantage of emerging cloud computing technique, processing unit can use big medical dataset stored in cloud platform to train naive Bayesian classifier, and then apply the classifier for disease diagnosis without compromising the privacy of data provider. In addition, the patient can securely retrieve the top-k diagnosis results according to their own preference in our system. Since all the data are processed in the encrypted form, our system can achieve patient-centric diagnose result retrieval in privacy preserving way waiting time.

VI. ACKNOWLEDGMENTS

For everything we get, the credit goes to all those who had helped us to complete this survey successfully. I am thankful to " Dr. V. S.Bidvefor guidance and review of this paper. I would also like to thanks, all faculty members of SKN College of Engineering".

REFERENCES

[1] H. Monkaresi, R. A. Calvo, and H. Yan, "A machine learning approach to improve contactless heart rate monitoring using a webcam", IEEE J. Biomedical and Health Informatics, vol. 18, no. 4, pp. 1153–1160, 2014.
[2] R. Lu, H. Zhu, X. Liu, J. K. Liu, and J. Shao, "Toward efficient and privacy-preserving computing in big data era," IEEE Network, vol. 28, no. 4, pp. 46–50, 2014.
[3] Y. Tong, J. Sun, S. S. M. Chow, and P. Li, "Cloud-assisted mobile-access of health data with privacy and auditability," IEEE J. Biomedical and Health Informatics, vol. 18, no. 2, pp. 419–429, 2014.
[4] Y. Elmehdwi, B. K. Samanthula, and W. Jiang, "Secure k-nearest neighbor query over encrypted data in outsourced environments," arXiv preprint arXiv:1307.4824, 2013.
[5] Y. Rahulamathavan, S. Veluru, R. Phan, J. Chambers, and M. Rajarajan, "Privacy-preserving clinical decision support system using Gaussian kernel based classification," IEEE Journal of Biomedical and Health Informatics, pp. 56–66, 2014.
[6] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with na¨ıvebayes," Expert Systems with Applications, vol. 36, no. 3, pp. 5432–5435, 2009.
[7] R. Bellazzi and B. Zupan, "Predictive data mining in clinical medicine: current issues and guidelines," International journal of medical informatics, vol. 77, no. 2, pp. 81–97, 2008.
[8] B. K. Samanthula, Y. Elmehdwi, and W. Jiang, "k-nearest neighbor classification over semantically secure encrypted relational data," arXiv preprint arXiv:1403.5001, 2014.
[9] X. Yi and Y. Zhang, "Privacy-preserving naive bayes classification on distributed data via semi-trusted mixers," Information Systems, vol. 34, no. 3, pp. 371–380, 2009.
[10] T. T. A. Dinh and A. Datta, "Stream on the sky: Outsourcing access control enforcement for stream data to the cloud," arXiv preprint arXiv:1210.0660, 2012.
[11] Wyatt JC, Liu JL. J Epidemiol, "Basic concepts in medical informatics", Community Health 2002; 56(11): 808-812.
[12] Barnett GO, Cimino JJ, Hupp JA, Hoffer EP, "DXplain. An evolving diagnostic decision-support system", July 3, 1987.