Comparative analysis using Hungarian dataset

The comparative analysis of the medical data classification methods using the Hungarian dataset is depicted in figure 5. Figure 5.a shows the accuracy of methods for various training percentages based on the Hungarian dataset. With 70% training, the accuracy of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5, 0.583, 0.625, 0.5625,0.8383, and 0.8543, respectively. With 80% training, the accuracy of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5, 0.5, 0.6875, 0.3125,0.8571, and 0.8638, respectively. Figure 5.b depicts the fitness of the methods for various training percentages based on the Hungarian dataset. When the training percentage is 70, the fitness of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.4444, 0.583, 0.6528, 0.5632, 0.8656, and 0.8762, respectively. When the training percentage is 80, the fitness of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.4455, 0.5, 0.708, 0.3105, 0.8799, and 0.8881, respectively.

Figure 5.c depicts the sensitivity of the methods for various training percentages based on the Hungarian dataset. When the training percentage is 70, the sensitivity of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5833, 0.8322, 0.5833, 0.5714, 0.9583, and 0.9696, respectively. When the training percentage is 80, the sensitivity of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.6364, 0.6662, 0.6364, 0.3333, 0.9751, and 0.9819, respectively. Figure 5.d depicts the specificity of the methods for various training percentages based on the Hungarian dataset. When the training percentage is 70, the specificity of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.25, 0.3339, 0.75, 0.5556,0.80, and 0.8046, respectively. When the training percentage is 80, the specificity of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.2, 0.3338, 0.8, 0.2857,0.8074, and 0.8186, respectively.

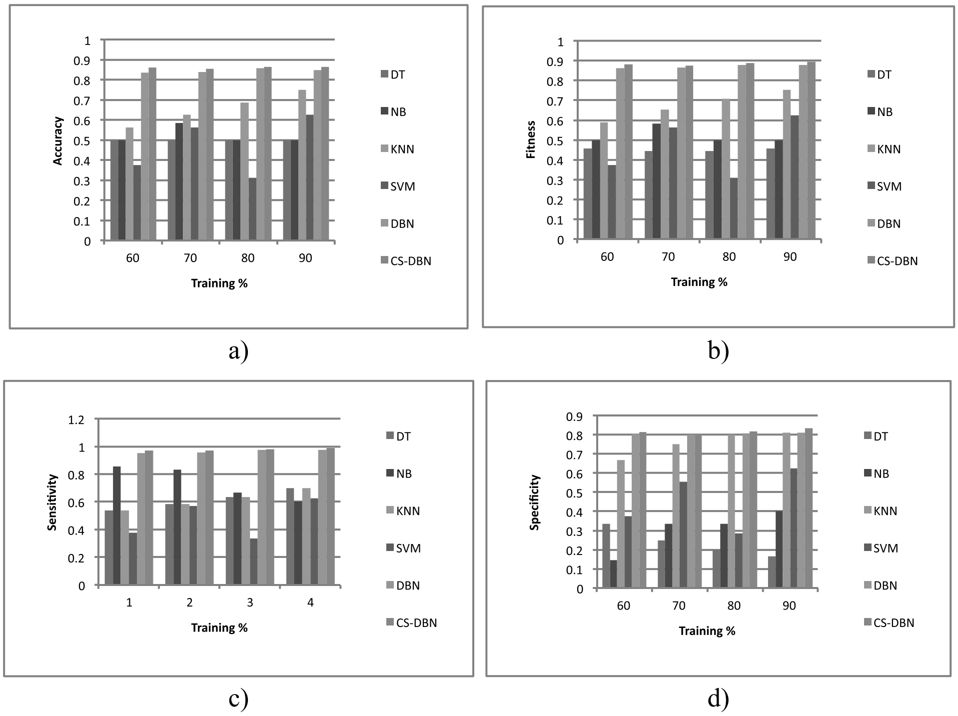

Figure 4. Analysis using Hungarian dataset based on a) accuracy, b) fitness, c) sensitivity, d) specificity.

Comparative analysis using Switzerland dataset:

The comparative analysis of the medical data classification methods using the Switzerland dataset is depicted in figure 6. Figure 6.a shows the accuracy of methods for various training percentages based on the Switzerland dataset. With 70% training, the accuracy of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5, 0.583, 0.625, 0.6875, 0.8441,and 0.863, respectively. With 80% training, the accuracy of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5, 0.5, 0.6875, 0.625, 0.8412, and 0.8623, respectively. Figure 6.b depicts the fitness of the methods for various training percentages based on Switzerland dataset. When the training percentage is 70, the fitness of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.4444, 0.583, 0.6528, 0.6895,0.8725, and 0.8858, respectively. When the training percentage is 80, the fitness of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.4455, 0.5, 0.708, 0.6306, 0.8726, and 0.8882, respectively.

Figure 6.c depicts the sensitivity of the methods for various training percentages based on Switzerland dataset. When the training percentage is 70, the sensitivity of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.5833, 0.8322, 0.5833, 0.7143, 0.9653,and 0.9829, respectively. When the training percentage is 80, the sensitivity of DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN is 0.6364, 0.6662, 0.6364, 0.6667, 0.976,and 0.9954, respectively. Figure 6.d depicts the specificity of the methods for various training percentages based on Switzerland dataset. When the training percentage is 70, the specificity of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.25, 0.3339, 0.75, 0.6667, 0.8023, and 0.8115, respectively. When the training percentage is 80, the specificity of the methods, such as DT, NB, K-NN, SVM, DBN, and the proposed CS-DBN, is 0.2, 0.3338, 0.8, 0.6, 0.8005, and 0.8069, respectively.

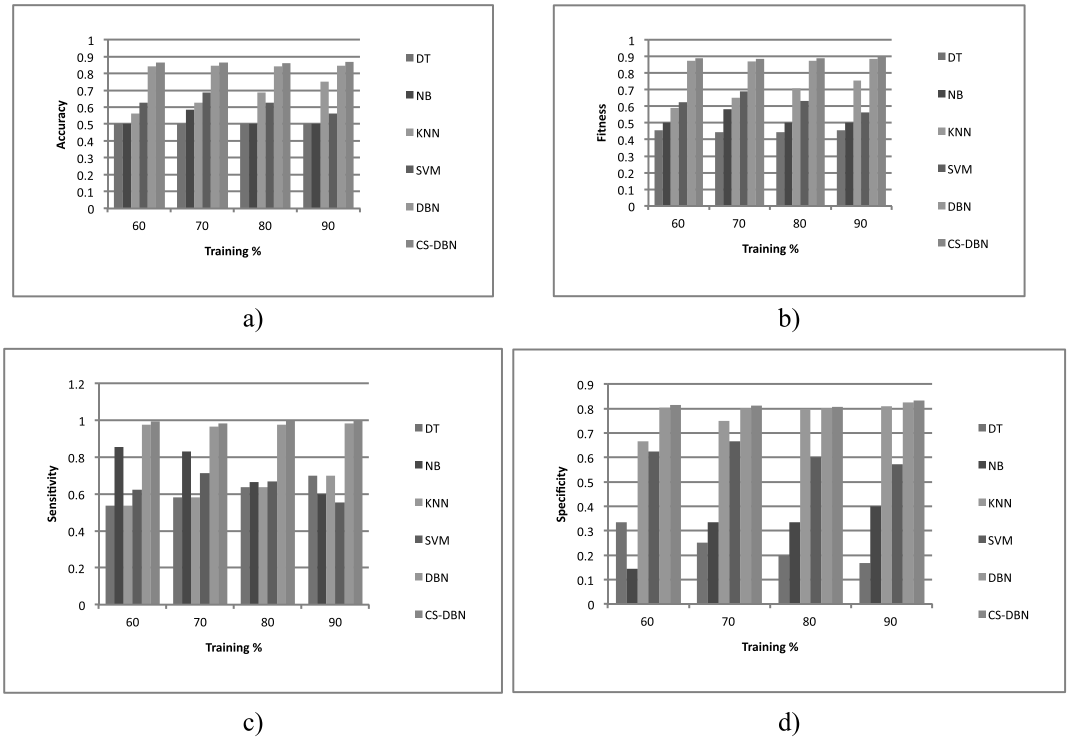

Figure 5. Analysis using Switzerland dataset based on a) accuracy, b) fitness, c) sensitivity, d) specificity.

Analysis based on the computational time

The analysis of the proposed method with the existing methods in terms of computational time is provided in this section. Table 2 depicts the computational time of the proposed method and the existing methods, such as DT, NB, K-NN, SVM, and DBN, in which the proposed system has less computation time of 6 sec.

Table 2. Computational time.

Methods	DT	NB	K-NN	SVM	DBN	Proposed CS-DBN
Time (Sec)	13	11.5	10.4	8	7.5	6

CONCLUSION

Preserving the privacy of medical data in the ontology-based systems is a critical need, especially in the case when the system is used by more numbers of users with various privileges and is distributed over applications. Thus, it is necessary to take steps for the preservation of the medical data of the patients. This paper aims to preserve confidential medical data with the introduction of a medical data classification method. The proposed CS-DBN method works based on three main steps, namely, generation of privacy preserved data, construction of ontology, and classification. The deep convolutional kernel approach is utilized for the provision of data confidentiality with the generation of optimal coefficients. The ontology is developed with the terms related to cardiac heart disease for classification. The classification is carried out using deep belief network (DBN) that is trained by crow search algorithm (CSA). The analysis of CS-DBN is performed in terms of the metrics, namely, fitness, accuracy, sensitivity, and specificity, and it produces the higher fitness, accuracy, sensitivity, and specificity of 0.9007, 0.8842, 1, and 0.8408, respectively. In future, the data classification will be based on any hybrid optimizations, and the analysis will be done using more medical databases.

REFERENCES

Alabdulkarim, A., Al-Rodhaan, M., Ma, T & Tian, Y. (2019). PPSDT: A Novel Privacy-Preserving Single Decision Tree Algorithm for Clinical Decision-Support Systems Using IoT Devices, **3**(19): 1.

Al-Aidaroo, K.M., Bakar, A.A & Othman, Z. (2012). Medical Data Classification with Naive Bayes Approach, Information Technology Journal, **11**(9): 1166-1174.

Arul. V.H, V.G. Sivakumar, Ramalatha Marimuthu & Basabi Chakraborty. (2019). An Approach for Speech Enhancement Using Deep Convolutional Neural Network, Multimedia Research (MR), **2**(1): 37-44.

Askarzadeh, A. (2016). A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm, Computers & Structures, **169**: 1-12.

Bernabe, J.B., Perez, G.M., Antonio, F & Gomez, S. (2015). Intercloud Trust and Security Decision Support System: an Ontology-based Approach, Journals on Grid Computing.

Biskup, J & Bonatti, P.A. (2004). Controlled query evaluation for enforcing confidentiality in complete information systems, International Journal of Information Security, **3**(1)14-27.

Biskup, J & Weibert, T. (2008). Keeping secrets in incomplete databases, International Journal of Information Security, **7**(3): 199-217.

Cachin, C & Haas, R. (2010). Dependable Storage in the Intercloud, IBM Research Report RZ 3783.

Dasarathy, B.V. (1980). Nosing aroung the neighbourhood: A new system structure and classification rule for recognition in partially exposed environments, IEEE Transactions on Pattern Analysis and Machine Intelligence, **2**: 67-71.

Duba, R.O & Hart, P.E. (1973). Pattern Classification and Scene Analysis, New York: Wily.

Fan, W., Chan, C.Y. & Garofalakis, M.N. (2004). Secure xml querying with security views, In Proceedings of SIGMOD, 587-598.

Farooq, K & Hussain, A. (2016). A novel ontology and machine learning driven hybrid cardiovascular clinical prognosis as a complex adaptive clinical system,Complex adaptive syatems modelling, **4**(12).

Geibel, P., Trautwein, M., Erdur, H., Zimmermann, L., Jegzentis, K., Bengner, M., Nolte, C.H & Tolxdorff, T. (2015). Ontology-Based Information Extraction: Identifying Eligible Patients for Clinical Trials in Neurology, Journal on Data Semantics, **4**(2): 133-147.

Grau, B.C. (2010). Privacy in Ontology-based Information Systems: A Pending Matter, Semantic Web, **1**(1,2): 137-141,.

Heart disease dataset, “https://archive.ics.uci.edu/ml/datasets/Heart+Disease” accessed on January 2019.

Karlekar, N.P & Gomathi, N. (2017). Kronecker product and bat algorithm-based coefficient generation for privacy protection on cloud, International Journal of Modeling, Simulation and Scientific Computing, **8**(3).

Karlekar, N.P & Gomathi, N. (2018). OW-SVM: Ontology and whale optimization-based support vector machine for privacy-preserved medical data classification in cloud, **31**(12).

Keller, J.M., Gray, M.R & Givens, J.A. (1985). A Fussy-K-Nearest Neighbor Algorithm, IEEE Transactions on System, Man, and Cybernetics, **15**(4): 580-585.

Levy, A.Y. (1996). Obtaining complete answers from incomplete databases, In Proceedings of Very Large Data Bases, 402-412.

Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data, IEEE Transactions on Parallel Distribution Systems, **27**(2): 340-352.

Pramod P Jadhav & SD Joshi, ACADF(2019): Ant Colony Unified with Adaptive Dragonfly Algorithm Enabled with Fitness Function for Model Transformation, Springer, Singapore:101-109.

Ram, C.P & Sreenivaasan, G. (2010). Security as a Service (SasS): Securing user data by coprocessor and distributing the data, Trendz in Information Sciences & Computing (TISC2010), 152-155.

Stouppa, P & Studer, T. (2007). A formal model of data privacy, In Proceedings of Perspectives of System Informatics 06, 4378.

Tao, M., Zuo, J., Liu, Z., Castiglione, A & Palmieri, F. (2018). Multi-layer cloud architectural model and ontology-based security service framework for IoT-based smart homes, Future Generation Computer Systems, **78**(3): 1040-1051.

Vito Racanelli, Claudia Brunetti, Valli De Re, Laura Caggiari, Mariangela De Zorzi, Patrizia Leone, Federico Perosa, Angelo Vacca & Franco Dammacco, (2011). Antibody Vh repertoire differences between resolving and chronically evolving hepatitis C virus infections, Public Library of Science, **6**(9).

Vojt, B.J. (2016). Deep neural networks and their implementation, Department of Theoretical Computer Science and Mathematical Logic, Prague.

Whitney, A & Dwyer, S.J. (1966). Performance and implementation of K-nearest neighbour decision rule with incorrectly identified training samples, In Proceedings of 4th Annual Allerton of Conference On Circuits Band System Theory.

Zardari, M.A., Jung, L.T. & Zakaria, N. (2014). K-NN Classifier for Data Confidentiality in Cloud Computing, International Conference on Computer and Information Sciences (ICCOINS).

Zhou, L., Zhang, C., Karimi, I.A & Kraft, M. (2018). An ontology framework towards decentralized information management for eco-industrial parks, Computers & Chemical Engineering, **118**: 49-63.