ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 11, pp.41659-41663, November, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

CONCEPTS OF DENTAL INFORMATICS AND MACHINE LEARNING IN ORAL HEALTH CARE

Shruthi Nambiar, K., *Dr. Shwetha Nambiar, K. and Dr. Supriya Bhat

Multimodal Perception Lab, International Institute of Information Technology, Bangalore-560100

ARTICLE INFO

Article History:
Received 19th August, 2016
Received in revised form
22nd September, 2016
Accepted 07th October, 2016
Published online 30th November, 2016

Key words:

Dental informatics,
Data mining,
Machine learning,
Evidence based dentistry,
Oral cancer.

ABSTRACT

Dental informatics is a sub-discipline of biomedical informatics and is the implementation of information science to improve practice of dentistry, research, education and management. This is an interdisciplinary field which applies knowledge and techniques from other disciplines such as information science, computer science, cognitive science and telecommunications. Image processing, digital imaging, computerized patient history records, clinical decision support and teledentistry are some popular research topics in dental informatics. It is a new field that supports patient care bridging the gap between researchers and clinicians. This field utilizes information science and it's applications to help practitioners utilize recent advances and discuss about cases with diagnostic dilemma (evidence based dentistry). The advances made in the field of machine learning and data mining has equipped researchers with newfound approaches to tackling traditional challenges like diagnosis and risk assessment. If a sufficiently large dataset is available to train the machine learning models, then we can build a fairly accurate classifier. The collection of this data is one of the initial challenges and hence maintaining electronic oral health records, pooling of genetic data globally will help create a large database that can be used to effectively predict the diagnosis and prognosis of specific diseases. However awareness and usage of these alternate approaches is low due to lack of relevant insights among the practicing dental community. This article aims to bridge this gap and highlight the concepts of dental informatics, machine learning and its applications in oral health care.

Citation: Shruthi Nambiar, Dr. Shwetha Nambiar, K. and Dr. Supriya Bhat, 2016. "Concepts of dental informatics and machine learning in oral health care", *International Journal of Current Research*, 8, (11), 41659-41663.

INTRODUCTION

Informatics is a word derived from the term "Information Science", which implies collection, storage, classification, retrieval and dissemination of data. (Amin *et al*., 2009; Chhabra *et al*., 2016) Dental informatics is a new field that is currently emerging which supports clinical health care, involves new insights in research, knowledge and management of oral diseases. It is the application of information science and its various disciplines, including cognitive science, genetic engineering, human psychology science, decision graphs, mathematical and human factors engineering. (Schleyer and Spallek, 2000; Kirshner, 2003) Knowledge of dental informatics can enlighten them with the new trends and innovations in practices, discuss rare cases in public platform and over all uplift the state of art in education, research and contribute to patient welfare. (Schleyer and Spallek, 2000) Dentists must be self motivated to learn all the modern developments to make informed flawless decisions in patient care. Unfortunately, the large spread dental community, especially in the third world countries are unacquainted with these new concepts and fail to understand the objectives, achievements and goals of dental informatics and how to impart it in their clinical practice. (Chhabra *et al*., 2016)

Corresponding author: Shruthi Nambiar, K.,
Multimodal Perception Lab, International Institute of Information Technology, Bangalore-560100.

Friedman defines four categories of informatics research (Four building blocks of informatics):

(1) Model formulation - creating risk assessment models standardized vocabularies, taxonomies and ontologies, that function as problem solving devices. This category forms frameworks using prior statistics obtained from patient health records.

(2) System development - designing an easy, usable system that will integrate the needs of clinician and patients. This category concentrates on the designing and integration of computer applications. Research domains include data science, decision science, human factors engineering, telecommunications and artificial intelligence helping in practice administration, decision making and treatment planning.

(3) System installation - implementing, adopting and sustaining systems by understanding the psychological and cultural identities of the people and the various organizations

that are affected by this system. This category addresses the issues associated with the collection, systematic organization and execution of risk assessment models. This research field also includes organizational and management theory, cognitive psychology and human computer interaction.

(4) Study of effects - considers the system results on a specific individual's health, groups and population based levels. This category examines the processes and outcomes resulting from the use of these applications. This research uses both, qualitative and quantitative methods that are established to study the effects. (Kirshner, 2003)

The goals of dental informatics

1. Develop and provide competent dental care delivery system
2. To help the clinician achieve adequate or improved patient care
3. Solve patient treatment related problems through evidence (EBD- evidence based dentistry)
4. Competency between clinicians and researchers to be maintained, to implement new innovations in clinical practice. (Dadu *et al*., 2013)

Fundamentals of dental informatics

1. Interoperability in dental information system

Ability to exchange data and communicate effectively in various systems within information technology.

2. Electronic prescription of medicines

This approach can reduce the medication errors ensuring patient safety. Dispensing and monitoring of pharmacological therapies, new insights into immunological therapies and customized medication can be aimed with this advanced approach. Electronic Oral Health Records (EOHR) is one of the forth coming advances in the field of dental informatics. It can act as a central repository of dental information, establishing a positive patient and dentist relationship. Records will contain all details of patient's oral condition – clinical details, necessary radiographs, psychological and medical information. This will improve the efficacy to formulate accurate treatment plan and prevent avoidable errors with these clinical decision support tools.

3. Retrieval of available information

Data mining helps in searching disease information and relevant documents by exploring metadata like standalone databases type and hypertext network databases like internet. This helps in gathering all information from the data pool about a specific disease, thus a better understanding of the disease process.

4. Extraction of the retrieved information

Extraction of the relevant information necessary to design a specific treatment plan.

5. Visualisation of the extracted information

Computer supported abstract data that can be visualized to analyse a specific case.

6. Artificial intelligence

Disease prediction models are designed to flawlessly diagnose a case and approach the therapeutic measures.

7. Laboratory information system

Programs are designed such that it can automatically interpret the test results when relevant data are entered.

8. Clinical decision support system (CDSS)

These are active knowledge systems that gives access to patient data, generates case-specific advice for the safety of the patient and to improve the prognosis. (Balas *et al*., 1996; Shiffman *et al*., 1999; Jawahar Babu *et al*., 2012; Longo *et al*., 2005; Zhan *et al*., 2006; Aarts and Berg, 2006; Grimshaw and Russell, 1993)

Opportunities in dental informatics

1. Organized dental data management: Longitudinal, lifetime, comprehensive and patient-centered dental records
2. Data collection without revealing patient identity: Patient records that are universally accessible but ensuring patient privacy and confidentiality
3. Significant reduction of ambiguity in diagnosis and time management (in emergency cases)
4. Treatment planning entirely based on empirically determined best practices
5. Expansion in knowledge-based ontology of new dental concepts
6. Understanding evidence based learning of risk factors, diagnosis, prevention and therapeutic outcomes (Risk assessment models)
7. Comprehensive EOHR mounted flawlessly integrated into medical record
8. To develop and maintain an organized nationwide and worldwide oral health database
9. Follow up of patients
10. Mapping of human genetics and identifying the genetic factors involved in a particular disease and exploring genetic disparities among different ethnic groups. (Schleyer and Spallek, 2000; Sittig *et al*., 2003; Johnson, 2003)

Data mining technique

Data mining technique involves the use of sophisticated data analysis tools to discover previously not known, specific patterns and relationships in large set of data. These tools can include statistical records, mathematical algorithm and machine learning methods in early detection of a disease.This technique is employed to create a unique method of identification of a particular disease. It is essential to first collect and assemble the available data, integrate and then process through machine learning techniques. This developed information system, is a set of complex data, that can be used by both, physicians and the patients alike to know a person's disease status and severity. Maintaining of oral health records, pooling in of data in large databases help us to gain knowledge about the disease and its treatment. These days, Knowledge Discovery in Databases (KDD) has become a popular research tool for medical researchers. This database records and saves large

volumes of sensitive information which can be used to understand the disease process. It includes data mining techniques to identify and exploit certain patterns and relationships among large number of variables and help predict the outcome of a disease using the historical case records stored within these databases (Ramachandran *et al*., 2014; Kharya, 2012; Han and Kamber, 2001).

Machine learning

New technologies in machine learning (ML) like deep learning has given a ray of hope to biomedical research. Machine learning algorithms are computer programs that try to predict either the behavior, type, incidence of the tumour/ disease based on past results. The eventual goal of machine learning in diagnosis is to have a trained machine learning algorithm that can accurately predict the type and severity of the disease based on the gene expression levels or other specific data from the patient (Kourou *et al*., 2015; Kharya *et al*., 2013). This aids the physician in accurate diagnosis, predicting a disease outcome and plan treatment. Factors such as family history, age, food habits, weight, deleterious oral habits, clinical variables, exposure to certain carcinogens, cellular parameters, histological parameters and molecular biomarkers play an important role in disease prediction. ML techniques use different techniques, algorithms to identify patterns and relationships, between various factors causing disease, from a complex database. This creates a risk assessment model which can classify the disease, predict the susceptibility, recurrence and survival rate in that particular disease. (Cruz and Wishart, 2006; Hsieh *et al*., 2012) Disease assessment and prediction models like Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs) and Decision Trees (DTs) have been used currently in medicinal research for effective and accurate decision making. The Cancer Genome Atlas Research Network (TCGA) is a community resource project, that is currently providing large scale genomic data through high-throughput genome technologies about specific tumor types to design customized medicines. The prediction model is analyzed based on sensitivity, susceptibility and accuracy. (Hsieh *et al*., 2012; Tomczak *et al*., 2015) Large amount of initial labeled data are split and stored into subsets: (1) Holdout Method, (2) Random Sampling, (3) Cross-Validation and (4) Bootstrap. In the Holdout method, the data samples are classified as training and the test sets. A classification model is created from the training set and it's efficacy is checked on the test set. Random sampling is similar to Holdout method but checked several times, randomly choosing the training and test instances. In cross validation approach, original data set is successfully covered as every sample is used once for testing but same number of times for training. In bootstrap approach, the separated samples are divided into training and test sets, placed back into the database after they are trained. (Kourou *et al*., 2015; Kharya *et al*., 2013)

Informatics and machine learning in oral cancer

The main aim of oral cancer prediction and prognosis are distinct and differ from the methods of cancer detection and diagnosis. The three main factors considered for cancer prediction/prognosis are: 1) cancer susceptibility prediction (i.e. assessing the various risk factors), 2) cancer recurrence prediction and 3) survival rate prediction. In the first factor, various carcinogens and risk factors are considered to predict the development of cancer prior to the disease occurence. In the second factor, cancer recurrence prediction is done after resolution of primary tumour. In the third factor, prognosis is predicted (survival rate, metastatic potential, drug sensitivity). The cancer recurrence rate, survival rate prediction is entirely based on adequate diagnosis. (Ahmad *et al*., 2013) In India, most oral cancer cases report at stage III or stage IV where the prognosis is bad. This is mainly due to ignorance of the patients and inability to diagnose accurately by rural doctors. Access to worldwide information on oral cancer would probably help target oral cancer at an early stage and prevent metastasis and other complications. (Nambiar and Hegde, 2016; Nambiar *et al*., 2016) Oral cancer diagnosis and prediction of prognosis typically involves multiple physicians from different specialties (Oral medicine and radiology, Oral and maxillofacial surgery, Oral Pathology and microbiology). Based on multiple clinical factors, including the age and general health of the patient, the location and type of cancer, the grade and size of the tumor, the initial diagnosis is done. Clinical (patient-based) and demographic (population-based) information must all be carefully integrated by the attending physician. Histological (cell-based) analysis along with different subsets of biomarkers may be used to come up with a reasonable prognosis. (Cruz and Wishart, 2006)

With current progress in understanding a disease at a molecular level, genomic analysis (DNA sequencing, RNA sequencing, microarrays), proteomic analysis (protein chips, tissue arrays) and imaging (MRI, PET, micro-CT) technologies, large amount of molecular-scale information about patients or tumors can be readily acquired. Molecular biomarkers, such as somatic mutations in certain genes (p53, Ki-67, Bax, Bcl, Survivin, C-Jun), the appearance or expression of certain tumor proteins (upregulation/down regulation) or the chemical environment (acidic/alkaline) of the tumor serve as very powerful prognostic/predictive indicators. More recently, combinations or patterns of multiple molecular biomarkers have been found to be even more predictive than single component tests. If these molecular patterns are combined with macro-scale clinical data (tumor type, hereditary aspects, risk factors), the robustness and accuracy of cancer prognosis and predictions will enhance over time. (Kourou *et al*., 2015; Cruz and Wishart, 2006; Meldrum *et al*., 2011)

Current challenges in dental informatics

1. Universal access to computers and high-speed internet connections.
2. National list of authorized dental practitioners
3. Collections of large databases of patient information
4. Lack of literature on dental informatics.
5. Lack of familiarity of dental fraternity in use of informatics (Chhabra *et al*., 2016; Kirshner, 2003; Dental Informatics: A Work in Progress, 2003; Sittig *et al*., 2003; Spallek *et al*., 2007)

Further recommendations

Dental informatics is like an "encyclopedia for dental health and disease". For the easy accessibility of all the information to people worldwide and enhance its capability to bring about its research goals, some of the recommendations put forth are:

1. Training dental bioinformaticians worldwide.

2. Government bodies should fund research and well deserved candidates should be recruited
3. It should be a career oriented field and researchers should be getting enough career opportunities to earn their livelihood.
4. Dental informatics should be capable of addressing dental public health problems at an international level. (Patwardhan *et al*., 2015; Schleyer, 2003; Schleyer, 2003; Levine *et al*., 2008; Kirshner, 2003; Seldin, 2001)

Dental Informatics Internship Program

These programs already exist in western countries. Starting similar internship programs in India will bring awareness, open up new arenas of opportunities to learn and practice dental informatics. Proper education, training and support systems should be developed and incorporated in dental fraternity with good surveillance for best implementation in patient welfare. (Chhabra *et al*., 2016; Spallek *et al*., 2007)

Online research Community

E-communities should utilize the worldwide databases to overcome various obstacles faced by dental informaticians by bringing researchers and collaborators together at a resource hub. This will enhance the knowledge and abilities to face newchallenged in practice. (Chhabra *et al*., 2016; Spallek *et al*., 2007)

Conclusion

When the gap between the dentist and the dental informatics is bridged, it will pave a new path for innovations, decision making, real-time information diffusion and edification in the current practice. New research and international collaborations will open doors to enhanced quality of patient care. Integration of multidimensional heterogeneous data from large databases and multidisciplinary involvement and analysis can provide a promising tool for treating all diseases. The appropriate use of newer technology will reduce the financial burden and permit a better level of care with greater efficiency and productivity.

REFERENCES

Aarts J. and Berg M. 2006. Same systems, different outcomes – comparing the implementation of computerized physician order entry in two Dutch hospitals. *Methods Inf Med*., 45(1):53–61.

Ahmad LG, Eshlaghy AT, Poorebrahimi A, Ebrahimi M, Razavi AR. 2013. Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence. *J Health Med Inform*, 4:124.

Amin W, Kang HP, Egloff AM, *et al*. 2009. An informatics supported web-based data annotation and query tool to expedite translational research for head and neck malignancies. *BMC Cancer*, 9:396.

Balas EA, Austin SM, Mitchell JA, Ewigman BG, Bopp KD, Brown GD. 1996. The clinical value of computerized information services. A review of 98 randomized clinical trials. *Arch Fam Med*., 5(5):271–78.

Chhabra KG, Mulla SH, Deolia SG, Chhabra C, Singh J, Marwaha BS. 2016. Dental Informatics in India: Time to Embrace the Change. *Journal of Clinical and Diagnostic Research: JCDR*, 10(3):ZE12-ZE15.

Cruz JA. and Wishart DS. 2006. Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Inform*, 2: 59–77.

Dadu M, Aruna DS, Arora D, Arora R. 2013. Evolution of dental informatics: A cornerstone of dental practice. *IJDE*, 6:33-39.

Dental Informatics: A Work in Progress *ADR* December 1, 2003 17: 9-15.

Grimshaw JM. and Russell IT. 1993. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*, 342(8883):1317–22.

Han J. and Kamber M. 2001. Data Mining: Concepts and Techniques, 2nd ed., San Francisco, Morgan Kauffmann Publishers.

Hsieh N, Hung L, Shih C. 2012. Intelligent post operative morbidity prediction of heart disease using artificial intelligence techniques. *J Med Syst*, 36:1809-1820.

Jawahar Babu KVSN, Harshavardhan V, Kumar JSA. 2012. The role of information retrieval in knowledge Management. *International Journal of Social Science & Interdisciplinary Research*, 1(10):212–26.

Johnson LA. 2003. Biomedical Informatics Training for Dental Researchers. *Adv Dent Res*., 17:29–33.

Kharya S, Dubey D, Soni S. 2013. Predictive machine learning techniques for breast cancer detection. *IJCSIT*, 4:1023-1028.

Kharya S. 2012. Using data mining techniques for diagnosis and prognosis of cancer disease. *IJCSEIT*, 2:55-66.

Kirshner M. 2003. The Role of Information Technology and Informatics Research in the Dentist-Patient Relationship. *Adv Dent Res*., 17:77-81.

Kirshner M. 2003. The Role of Information Technology and Informatics Research in the Dentist-Patient Relationship. *Adv Dent Res*., 17:77-81.

Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. 2015. Machine learning applications in cancer prognosis and prediction. *CSBJ*, 13:8-17.

Levine AE, Bebermeyer RD, Chen JW, Davis D, Harty C. 2008. Development of an interdisciplinary course in information resources and evidence-based dentistry. *JDE*, 72:1067-76.

Longo DR, Hewett JE, Ge B, Schubert S. 2005. The long road to patient safety: a status report on patient safety systems. *J Am Dent Assoc*., 294(22):2858–65.

Meldrum C, Doyle MA, Tothill RW. 2011. Next-Generation Sequencing for Cancer Diagnostics: a Practical Perspective. *Clin Biochem Reviews*, 32(4):177-195.

Nambiar S, Hegde V, Yadav N, Hallikeri K. 2016. Improvization of conventional cytology by centrifuged liquid-based cytology in oral exfoliative cytology specimen. *J Cytol*., 33:115-9.

Nambiar S, Hegde V. 2016. Apoptosis in oral epithelial dysplastic lesions and oral squamous cell carcinoma: A prognostic marker. *Indian J Pathol Microbiol*., 59:284-6.

Patwardhan N, Bhaskar DJ, Bumb SS, Agali C, Punia H. 2015. Dental informatics in planning an effective oral health information system: A review article. *TMU J Dent*., 2(1):12-16.

Ramachandran P, Girija N, Bhuvaneswari T. 2014. Early Detection and Prevention of Cancer using Data Mining Techniques. *IJCA*, 97:48-53.

Schleyer T, Corby P, Gregg AL. 2003. A Preliminary Analysis of the Dental Informatics Literature. *Adv. Dent. Res*., 17:20-24.

Schleyer T, Spallek H. 2000. Dental informatics. A cornerstone of dental practice. *J Am Dent Assoc*., 132(5):605-13.

Schleyer T. 2003. Dental Informatics: A Work in Progress. *Adv Dent Res*., 17:9-15.

Seldin LW. 2001. The future of dentistry: An overview of a new report. *J Am Dent Assoc*., 132:1667-77.

Shiffman RN, Liaw Y, Brandt CA, Corb GJ. 1999. Computer-based guideline implementation systems: A systematic review of functionality and effectiveness. *J Am Med Inform Assoc*., 6(2):104–14.

Sittig DF, Kirshner M, Maupomé G. 2003. Grand Challenges in Dental Informatics. *Adv Dent Res*., 17:16–19.

Spallek H, Irwin JY, Schleyer T, Butler BS, Weiss PM. 2007. Supporting the Emergence of Dental Informatics with an Online Community. *Int J Comput Dent*., 10(3):247–64.

Tomczak K, Czerwińska P, Wiznerowicz M. 2015. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol*., 19(1A):A68-A77.

Zhan C, Hicks RW, Blanchette CM, Keyes MA, Cousins DD. 2006. Potential benefits and problems with computerized prescriber order entry: analysis of a voluntary medication error-reporting database. *Am J Health Syst Pharm*., 63(4):353–58.

****