CONCLUSION AND FUTURE DIRECTION
CONCLUSION AND FUTURE DIRECTION

Clinical Informatics is an upcoming and challenging domain that holds potential for development of new informatics based techniques in the field of clinical science, especially in India. The field entails development of new Information systems, storage solutions and mining techniques for effective health care pertaining to a patient at the level of individual and population. In this research work, I have tried to address the applicability of knowledge discovery process in clinical informatics and plan for devising new strategies for future.

For the data storage problem lying in façade of clinical informatics, I have proposed a clinical dimensional model that can be used for development of clinical data mart. The model has been designed keeping in consideration temporal storage of patient's data with respect to all possible clinical parameters which includes qualitative/quantitative and also image based data. Availability of said data for each patient can be then used for application of data mining techniques for finding the correlation of all the parameters at the level of individual and population. A future direction of the proposed solution can be the large scale implementation of this prototype in form of a patient centric management system. It will provide a platform for longitudinal storage of patient's data in non-volatile, subject-oriented and integrated form, corresponding to each visit or any diagnostic process undergone into one of the networked hospital.

As an application of data mining, I performed a case study on dataset of brain tumor patients (primary stage). The association mining performed in this study suggests - high values of Creatinine, Blood Urea Nitrogen (BUN), SGOT & SGPT to be directly associated with tumor occurrence for patients in the primary stage with atleast 85% confidence and more than 50% support. Also, based on the parameters identified, I propose a normalized regression model along with Haemoglobin content, Alkaline Phosphatase and Serum Bilirubin for prediction of occurrence of STATE (brain tumor) as 0 (absent) or 1 (present). Parameter identified in this study along with the proposed predictive model is currently under observatory and validation of practicing physicians and oncologist across hospitals in India.

Temporal mining of clinical data is one of the less explored domain in this field. Pertaining to this challenging domain, I have proposed a novel "SN" algorithm, to map clinical parameters found to be associated with a disease, and to its state at various temporal points. The proposed algorithm is based on Jacobian's approach, which augurs the state of a disease 'S_n' at a given temporal point 'T_n' by mapping the derivatives with the temporal point 'T_0', whose state of disease 'S_0' is known, for estimating its Jacobian determinant. The proposed algorithm has been applied as
a case study on a temporal clinical data set of brain tumor patients. A very high prediction accuracy of $\sim 100\%$ is been observed for a brain tumor state $S_n$ for any temporal point $T_n$, provided. However, the algorithm needs to be further validated on other clinical data sets. Also the future need is to develop a generalized computational tool that can be used for auguring future state associated to any disease based on the clinical parameters found to be associated with it. The future state will be predicted based on clinical values observed at historical temporal points.

Another data mining study that I have performed in this work, is a cross-sectional study to identify cognitive analyzers among cognitive screening and clinical parameters for the low ($\leq 350m$) & highlander ($\geq 1500m$) population staying at higher altitude ($\geq 4300m$) for prolonged duration, that can be associated with cognitive impairment, beck depression inventory and insomnia. Visuospatial Executive, Attention, Coordination & Learning, Object recognition, Procedural Memory, Recall, Language for lowlander population while Procedural Memory, Coordination and Learning, Visuospatial Executive, Recall, Language for highlander population respectively, are the key MDCST parameters identified for analyzing the cognitive performance with observed p-value $\leq 0.05$. For low & highlanders respectively, different set of cognitive performance based associative rules have been observed at least with 30% support and more than 60% confidence for behavioural and clinical measures. This particular study was performed by the research support and data received from DIHAR, DRDO, India. The results observed in the study have been confirmed by the human cognition and physiology experts from DIHAR, DRDO, India. An interesting direction to work upon in this area would be to analyze the effect of parameters identified in this study longitudinally at the level of individual and population (lowlander and highlander) with respect to cognitive impairment, beck depression inventory and insomnia. Also it will be interesting to discover the key parameters associated to cognitive impairment, beck depression inventory and insomnia for the native population at an altitude of $\geq 4300m$.

The future holds lots of challenges and potential new developments in this field. Draft bill of EHR regulation in India is being proposed in parliament during 2013 winter session. Once the bill is passed as a regulation and there are definite policies associated for EHR implementation in India, would like to work in its development. Also would like to work on the possible scenario of integration of EHRs with Aadhar (unique identification system for Indian population). EHRs along with database based storage of clinical data further expands the possibilities regarding data mining thereby opening the door to a vast source of clinical data analysis, that would involve usage of existing as well as new algorithms.