Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 2, Number 5; April-June, 2015 pp. 411-415

http://www.krishisanskriti.org/acsit.html

A Reinforcement Learning Approach to E-Health

Kumari Aishwarya[1], S. Harshitha[2] and K. Chandrasekaran[3]

[1,2]Department of Computer Science and Engineering, NITK Surathkal Mangalore, India
E-mail: [1]k.aishwarya15@gmail.com, [2]kchnitk@ieee.org

Abstract—*Internet of Things generates massive amount of data everyday. It has been anticipated that by 2025 50 billion devices will be connected to the internet which would generate Zettabytes of data. It becomes very important to make sense out of this huge data. Hence, Machine learning plays an important role here to identify patterns and extract relevant data. eHealth is a revolutionising concept, which aims at redesigning health-care services and processes by the integration and use of communication among electronic devices at all levels. It aims in improvising global health by including patient involvement as much as possible in the decision making process.In this paper, we propose a framework that uses multi-agent reinforcement learning concept to enable hospital management system to learn and intelligently detect diseases quicker and suggest effective treatments to the patients, based on the global statistics, medical facilities, treatments etc across various regions. The framework also implement an effective epidemic alert system .This frame-work aids medical professionals to provide fast and effective treatments and prevent spread of diseases.*

Keywords: *E-Health, M-Health, IoT, MARL, MDP.*

1. INTRODUCTION

E-Health is better quality health services and a healthier life which is achieved through digital technology. It has the potential to empower patients, where each patient is well informed, has variety of choices and is more involved in the decision-making process .[8] The current E-health systems provide:

- Availability of patient health records among concerned professionals whenever and wherever needed.
- Access to prescribing options and giving prescriptions and sending them electronically to patients or pharma-cists.
- Providing treatments to the patients from a distance which includes 24-hours condition monitoring of pa-tients.
- Keeping track of best practice guidelines and epi-demics for developing effective treatments.
- Formation of Virtual healthcare teams which consists of healthcare professionals who co-ordinate, discuss and share information on patients through digital and electronic equipments[9].
- Mobile devices are used to collect patient level health data, provide health-care information to professionals, researchers, and patients, real-time monitor the pa-tients records and conditions.
- Huge amounts of heterogeneous data can be handled by high computing and data management capabilities.
- Software for automatically scheduling appointments, managing patients data and records and other admin-istrative tasks related to health.

One of the important features in E-Health includes profes-sionals getting access to secure and high quality information drawn from a variety of sources across the globe, so that they can offer more informed diagnoses and appropriate treatments which are time efficient and of good quality. High quality information includes details about various diseases, symptoms, causes, tests for detection, effective treatments, their side effects, preventive measures etc to patients grouped on basis of age, sex, other medical conditions etc. It involves using rich, related and important relevant data to improve health-care and develop new and better treatments. It should be able to collect the raw data from E-Health systems that real-time monitor patients to identify patterns of behaviour affecting patients. Once the patterns are identified, it should be able to recommend various suitable treatments for the patients.

Our proposed framework caters to the above feature. By applying Q-learning algorithm, a system is able to detect diseases and suggest treatments sooner and precisely, based on statistics that is calculated over huge amount of raw data. Using the concept of MARL each system in a region is able to communicate with the system of other regions and gain relevant information and update it regarding diseases, its treatments etc. It is also able to alert citizens of its region about a epidemic in neighbouring region and suggest preventive measures to them electronically to their PDAs such as smart phones. It also provides relevant details and treatment information about a disease which may be specific other regions, so that when such diseases are encountered the professionals have all necessary information readily available

2. MULTI-AGENT REINFORCEMENT LEARNING

2.1 Reinforcement learning

Reinforcement learning is an area of machine learning which takes inspiration from behaviorist psychology, it basi-cally involves how an agent takes actions in a given environ-ment, such that its overall reward is maximized. It requires Markov Decision Process (MDP) environment i.e., the next state of the agent is dependent solely on the current state and not on any of the previous states. The main difference between the traditional learning techniques and reinforcement learning is that reinforcement learning does not require the knowledge about the MDP and it targets large MDPs where exact methods become infeasible. Reinforcement learning aims to achieve a balance between exploration and exploitation by focusing on real-time performance. It is very important to make clever exploration mechanisms to ensure good performance of reinforcement learning algorithm. This is achieved from methods like epsilon greedy method.[4]

2.2 Q-Learning

Q-learning is a type of reinforcement learning technique that is model free. It aims to find an optimal action-selection policy for a given state in any given finite MDP. In the end, the expected reward of taking a given action in a given state is given by learning the function of action-value and thereafter the optimal policy is followed. A policy is a rule which the agent selects a particular action, given the state it is in. When such an action-reward function is learnt for a state, the optimised policy is built by simply selecting the action which yields the highest reward for each state. Q-learning handles problems with stochastic transitions and its rewards, without adopting to changes being required. It is proven that for any finite Markov Decision Process, Q-learning will, over the time find an optimised policy, which means that the expected value of the total rewards returns over all successive steps, which starts from the current state, is the maximum that can be achieved.[4]

The algorithm:

The MDP consists of:

1. Agent C
2. Finite number of States S.
3. Set of actions per state A.
4. Event E

By performing an action on a state for a particular event, the agent moves from one state to another. An agents acquires a reward by executing an action in a specific state.

The aim of the agent is to maximize its total reward. The agent does this by learning which action is optimal for each state, in the sense, of the expected values of the total rewards over all the future steps which starts from the current state, not only the immediate reward which results from an action-state pair.

The algorithm hence calculates the Quality of a state-action function:

$$Q : SXA \;\; > R \tag{1}$$

The new Q-value for present state is calculated using the equation:

$$Q[k] := Q[k \;\; 1] + (r[k] \;\; Q[k \;\; 1])[4] \tag{2}$$

Although most work on reinforcement learning has focused exclusively on single agents we can extend reinforcement learning to multiple agents if they are all independent. They together will outperform any single agent due to the fact

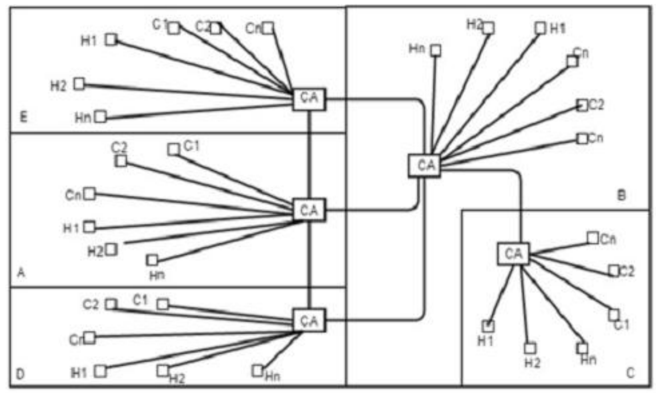

Fig. 1: Division of a region into districts

that they have more number of resources, more learning experience and a better chance of receiving higher rewards. There are three ways in which cooperation is achieved. First, agents communicate information such as events, actions or rewards. Second, agents communicate sequences of events, action, rewards which are called as episodes experienced by them. Third, agents communicate the learned decision policies.

3. METHODOLOGY

We use the concept of MARL with certain changes to design our framework. We divide each region into several districts. We install a central agent in each district, which has high computational capability and can store large amount of data. All the hospitals and clinics access the statistics and information stored in this central agent of its district to detect diseases and provide treatment, as well as update the information back on the central agent whenever a patient is treated. The central agents will then communicate their information and learning experience and increase the intelligence of the system as a whole.

3.1 The Framework

In the framework, we divide a region into n districts.(eg: A, B, C, D, E as shown in Fig. 1)

Each district maintains a central agent CA which is con-nected to E-health systems installed in hospitals (H1, H2,.. Hn) and clinics (C1, C2,... Cn) of that district. All these central agents are connected to each other across districts to enable MARL. Each District maintains a list of its adjacent neighbours. (Fig. 1)

3.2 Characteristics of the diseases

Each disease is characterized by common symptoms and specific symptoms. The probability of the disease being X, given that specific symptom for the disease X is observed in the patient, is higher than the probability that is calculated for a disease using common symptoms observed in a patient. Hence the system always checks for specific symptoms first before considering common symptoms

Diseases are also classified on basis of districts they are present in or are specific to. Every central agent maintains two tables, one contains all the diseases that are present in its district and the other contains all those diseases that are present in other districts. For a disease to be classified as present in a locality, the number of cases and frequency of the cases in that locality must be higher than a threshold value, which may vary for different diseases based on the complexity of the treatment, how much it affects the health etc.

The central agents prioritize the tests for the diseases belonging to its locality on basis of decreasing probability of how much symptoms are related to a disease. If the disease does not belong to a central agents district, then priority for the test for that disease is assigned on basis of how likely the symptoms imply the disease such that we can ignore its location.

3.3 The Algorithm

The algorithm used to detect a disease follows q-learning. The steps involved are:

1) Symptoms of a patient are fed into the system in hospital.
2) The system in hospital accesses the probability statis-tics table from central agent.
3) With the help of this statistic table we calculate the probability of the patient suffering from a disease and incrementally reward the tests for specific diseases starting from 1 according to increasing probability. These form the initial inputs for Q-table where the set of actions are the tests taken to confirm the diseases.
4) Based on the Q-table the system suggests that the patient undergoes test A before test B, where Q-value for test A is greater than that of B.
5) If the test results are positive, three steps that follows are:

The action is rewarded (eg: +10) in the Q-table.

The Q-table is sent to statistic table for update. The system goes into treatment phase.

6) If the test results are negative, two steps that follows are: The action is rewarded (eg: -5) in the Q-table The next action/ test B is executed
7) Repeat steps 4 to 6 until disease is detected We follow a same algorithm to suggest effective treatments for a particular disease based on details of a user.

Fig. 2 shows the flowchart for the above algorithm.

After the update of the statistic table in a central agent for that region, a trigger is issued which causes the central agent of a certain region to send the updated table to its adjacent neighbors based on the list it maintained. This update in neighbour central agents may further cause a trigger to share update table with adjacent regions other than the region that provided the updated table. Sharing of tables also occur in regular intervals.

In case an epidemic occurs in a district, it records its details and sets the emergency factor high and alerts its neighbouring districts about the epidemic, which then sends the updated table with the epidemic record to adjacent neighbours with the emergency factor set low.

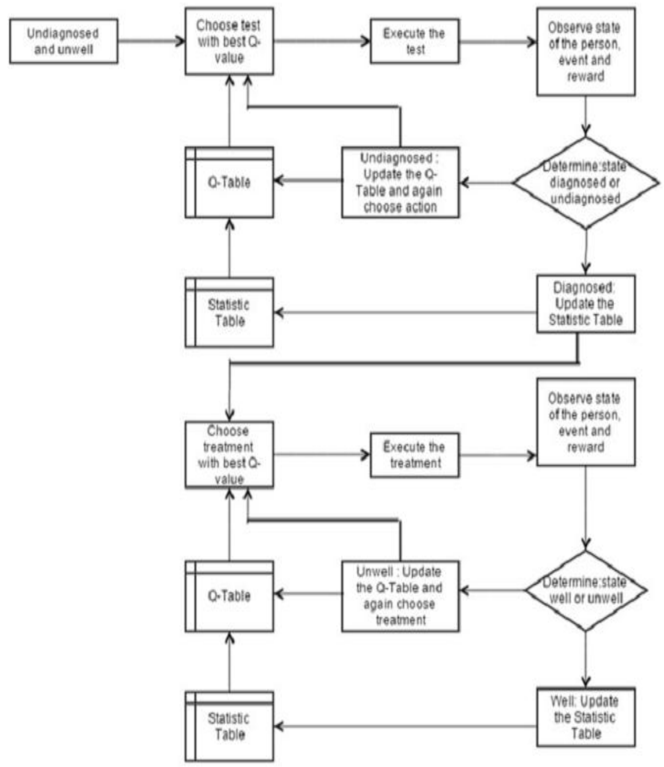

Fig. 2: The flowchart for Q-Learning

A district which has a high emergency factor will then alert its citizens by informing them electronically to their PDAs and create awareness about the disease, how it spreads and preventive measures to be taken.

Also an emergency factor is updated for a disease in a neighboring region only when it receives the updated informa-tion from the adjacent district where the epidemic is present

In this way all districts become aware of different dis-eases, its symptoms, treatment, emergency factor etc. Once the learning has reached stability, the systems start making optimized and accurate decisions. Hence, detecting diseases and suggesting effective treatments at a quicker rate than before.

3.4 Example Scenario

We consider the example case shown in Fig. 1. Table i, Table ii, Table iii and Table iv show the corresponding statistics each region would maintain with i and ii being common to all districts and iii being unique to C

Table 1: Statistic Table for Region C containing diseases belonging to its section

Statistic Table for Region C containing diseases belonging to its section								
	Common Symptoms				Total	Specific Symptoms		
	X	Y	Z	W		L	M	N
D1	70	10	60	10	150	130	0	0
D2	5	55	35	5	100	0	0	0
D3	75	60	15	50	200	0	180	0
D4	20	85	95	100	300	0	0	0
D5	120	10	10	110	250	0	0	270
Total	290	220	215	275	1000	130	180	270

Table 2: Statistic Table for Region C containing diseases belonging to A,B,D and E

Statistic Table for Region C containing diseases belonging to A,B,D and E									
	Common symptoms				Total	Specific Symptoms			Regions
	X	Y	Z	W		L	M	N	
D1	70	10	60	10	150	130	0	0	A,B,C,D,E
D2	5	55	35	5	100	0	0	0	A,B,C,D,E
D3	75	60	15	50	200	0	180	0	A,B,C,D,E
D4	20	85	95	100	300	0	0	0	A,B,C,D,E
Total	170	210	205	165	750	130	180	0	A,B,C,D,E

Table 3: Statistic Table for A,B,D and E containing diseases belonging to their districts

Statistic Table for A,B,D and E containing diseases belonging to their districts								
	Common symptoms				Total	Specific Symptoms		
	X	Y	Z	W		L	M	N
D1	70	10	60	10	150	130	0	0
D2	5	55	35	5	100	0	0	0
D3	75	60	15	50	200	0	180	0
D4	20	85	95	100	300	0	0	0
Total	170	210	205	165	750	130	180	0

Table 4: Statistic Table for Region A,B,D and E containing diseases belonging to other districts

Statistic Table for Region A,B,D and E containing diseases belonging to other districts									
	Common symptoms				Total	Specific Symptoms			Regions
	X	Y	Z	W		L	M	N	
D1	70	10	60	10	150	130	0	0	A,B,C,D,E
D2	5	55	35	5	100	0	0	0	A,B,C,D,E
D3	75	60	15	50	200	0	180	0	A,B,C,D,E
D4	20	85	95	100	300	0	0	0	A,B,C,D,E
D5	120	10	10	110	250	0	0	270	C
Total	290	220	215	175	1000	130	180	270	A,B,C,D,E

3.5 Example scenario

1) Initial Q-Values for each disease for Central Agent C for patient with symptoms X and W and suffering from D5:

   Probabilities:
   P(X,W—D1)=38.9,
   P(X,W—D2)=3.1,
   P(X,W—D3)=117.5,
   P(X,W—D4)=27.83,
   P(X,W—D5)=264.8
   Q values for testing for disease D(i) (i.e., action):
   D1: 3, D2: 1, D3: 4, D4: 2, D5: 5

   Patients disease is detected in one iteration

2) Initial Q-Values for each disease for Central Agents A,B,D,E for patient with symptoms X and W and suffering from D5:

   Probabilities: P(X,W—D1)=62.5, P(X,W—D2)=5.03, P(X,W—D3)=188.5, P(X,W—D4)=44.6, P(X,W—D5)=264.8 (obtained from Cs table)
   Q values for testing for disease D(i) (i.e., action):
   D1: 3, D2: 1, D3: 5, D4: 2, D5: 4

Since it is not likely for a person in regions other than C to have disease D5, but the symptoms are most related to D5, D5 is given priority up in the list but importance is also given to probability of other diseases present in current district. Hence, other diseases are also considered while finding the priority.

Here the patients disease is detected in two iterations.

4. SIGNIFICANCE OF THE PROPOSED MODEL

1) Predicting diseases at a faster rate. This also involves predicting those diseases which may not be specific to a particular district. This is achieved by the region-specific disease tables and disease specific symptoms feature. If a disease specific to other district is de-tected, the system will have available data on its details like treatment, tests etc. If it cant be treated in present region then the patient is treated in the disease specific region or given consultation on-line.
2) Suggesting effective treatments for a person based on his details and previous records. Effectiveness of a treatment varies according to a persons age, sex, genetic factors, on-going medication etc. The above framework suggests treatments in order of descending success rates for a person of specific details. Hence, improving the quality of health service.
3) Getting access to treatment information of a disease not handled before in that district and detecting it sooner.
4) Creating awareness and alerting people of a district about an epidemic/ disease that is likely going to occur in their locality because it has already occurred in its neighbouring district, prevents a great damage in health-care for that locality.
5) By precise detection of the disease it prevents the patient from going through unnecessary tests. Hence, reducing the cost for treatment.
6) Since every District maintains its central agent, de-tection of diseases more prone in that area is faster. It provides location based service.

5. CONCLUSION AND FUTURE WORK

The key contribution of this paper is how an unsupervised machine learning algorithm such as reinforcement learning will enable the devices in the E-Health system take smart decisions such as detecting the disease, suggesting effective treatment for a patient, alert patients about epidemic and in turn, improve their overall performance at run-time. The ability to seamlessly realize the self-learning techniques on top of the base machine learning context models without requiring further resources is the key for the success.

In the future, it should be able to detect the cause of the disease through pattern matching and identification. It should also be able to identify similar diseases so as to help in understanding them and develop a better treatment for them.

REFERENCES

[1] H. Kopka and P. W. Daly, A Guide to $\LaTeX$, 3rd ed. Harlow, England: Addison-Wesley, 1999.

[2] Kok-Lim AlvinYau, PeterKomisarczuk and PaulD.Teal, Reinforcement learning for context awareness and intelligence in wireless networks: Review, new features and open issues, Journal of Network and Computer Applications 35 (2012) , pp. 253267, January 2011.

[3] Arun kishore Ramakrishnan, Davy Preuveneers and Yolande Berbers,

[4] Enabling self-learning in dynamic and open IoT environments, 5th International Conference on Ambient Systems, Networks and Technologies(ANT-2014),Procedia Computer Science 32, pp.207214, 2014

[5] Ming Tan, Multi-Agent Reinforcement Learning: Independent vs. Coop-erative Agents, GTE Laboratories Incorporated, MIT

[6] Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An In-troduction, A Bradford Book ,The MIT Press Cambridge, Massachusetts London, England

[7] Avik Ghose, Priyanka Sinha, Chirabrata Bhaumik, Aniruddha Sinha, Amit Agrawal and Anirban Dutta Choudhary, UbiHeld - Ubiquitous Healthcare Monitoring System for Elderly and Chronic Patients, 1st Workshop on Human Factors and Activity Recognition in Healthcare, Wellness and Assisted Living, UbiComp13, September, 2013

[8] Micheal Vizard, IT Journalist , Machine Learning Comes to the Land of IoT, www.iotworld.com, January, 2014

[9] Stephen F. DeAngles, To Succed the internet of things will require ma-chines , Analytics and People, www.enterrasolutions.com, October,2014

[10] George Demiris, Editor, eHealth: Current Status and Future Trends, IOS Press, Amsterdam: 61. ISBN 1 58603 442 1

[11] Olga Anikeeva, Petra Bywood, EHealth Technologies in Primary Healthcare: Current Strengths and Limitations, Primary Health Care Research and Information Service, issue 21, November 2011.

[12] Jain, A. K. Data clustering: 50 years beyond k-means. Pattern Recog-nition Letters 31, 8 (2010), 651666.

[13] Cortes, C., and Vapnik, V. Support-vector networks. Machine learning 20, 3 (1995), 273297.

[14] Pal and C. Bhaumik. Low cost remote medical diagnosis and consulta-tion. 7th International Conference and Expo on Emerging Technologies for a Smarter World (CEWIT 2010).