

Machine learning and Systems Approaches in Relation of *Helicobacter pylori* To Ageing Diseases

D. Santosh Kumari^{1*}, DSVGK Kaladhar²

^{1,2}Dept. Of Microbiology and Bioinformatics, UTD, Bilaspur University, Bilaspur (CG), India

*Corresponding Author: dsantoshkumari87@gmail.com, Tel.: +91-9039842567

Available online at: www.ijcseonline.org

Abstract— *Helicobacter pylori* is one of the bacterial pathogen showing great effect and associations with ageing diseases. The complex data prediction was done with Weka software. When compared with data distribution between Males and Females in relation with class *H.pylori*, it is shown that females are more prone to ageing diseases (especially gastric and Obesity). Cluster Centroids method was predicted with more chances for people in Urban regions with family history suffers gastric, expiry due to heart stoke with gastric, drink tea, with tension and stress. AD Tree shown positive for *H.pylori* (N) showing stomach pain, family suffered previously and Age >29.5. Subset Evaluator method was shown that Stomach pain, Aging diseases and smoking are the attributes associated with *H.pylori*. The highest accuracy for the data was predicted with random Forest and K Star followed by J48, Bayes Net, Naive Bayes, LMT and Simple Cost.

Keywords— *Helicobacter pylori*, Ageing diseases, Systems approaches, Machine learning

I. INTRODUCTION

In the present decades, scientists and engineers are involved in developing machine techniques in the field of pattern recognition and predictions [1,2]. Classification and prediction methods improves the performance of learning is one of the most elementary aspect of intelligent behaviour [3]. The enormous amount of available data leads to new challenges to mine and discover hidden information [4,5].

Helicobacter pylori is a microaerophilic gastric pathogen belongs to gram negative bacterium, which resides in the duodenum and stomach causes tumours, ulcers, cancers like adenocarcinoma, gastric lymphoma and inflammation [6,7,8]. *H. Pylori* causes chronic, usually lifetime, inflammation of the gastric mucosa, that slowly progress to intestinal metaplasia and leads to health problems like cancer and heart stroke [9,10].

Prediction and analysis of the mechanisms involved in several aging diseases like kidney stones, Alzheimer's, Parkinson's, etc., are some of the challenging tasks [11]. Systems approaches are needed to study relationship between microorganisms like *Helicobacter pylori* with ageing diseases [12]. In the present study, a research was conducted based on several surveys regarding some of the relationships of aging diseases with *Helicobacter pylori*. Further advanced studies have to be conducted in these areas by the approaches of systems biology.

II. RELATED WORK

Helicobacter pylori are associated with rigorous forms of ageing and gastroduodenal disease. Several genes codes for antigens in *Helicobacter pylori* strains. Machine learning techniques are the most important existing techniques that can be applied to learn, predict and classify questionnaire data of *Helicobacter pylori* dataset. The present study focused on machine learning and protein interaction studies of *Helicobacter pylori* towards gastric and ageing diseases. In the present decades, Machine learning and Protein interaction techniques is a hybrid intelligent medical decision system used to predict the human data analysis and understanding mechanisms towards ageing diseases.

III. METHODOLOGY

The feedback data was collected from the patients of various hospitals in Bilaspur (CG) and Raipur regions of Chhattisgarh state (India). The data is entered into the database where prediction and evaluation of the data obtained based on the attributes given on the feedback form like Name, age, family history, gender, smoking, eating habits, exercise, stomach related problems, symptoms, etc. Nearly 25 to 30 attributes had been collected from different patients from various places. The data is predicted and analysis done by WEKA software. Based on prediction and evaluation protein network analysis in ageing diseases by String were performed for the system analysis of

Helicobacter pylori in relation with diseases (especially gastric and Obesity).

IV. RESULTS AND DISCUSSION

According to Data representation and evaluation, there are three main attributes stomach pain, family expired with Ageing disease and smoking which shows prevalence of *H.pylori* and are prone to aging diseases. When compared with data distribution between Males and Females in relation with class *H.pylori*, it is shown that females are more prone to gastric disease.

| Cluster centroids: | | | |
|--------------------|-------------------|-----------|-----------|
| Attribute | Full Data (60) | Cluster# | |
| | | 0 (27) | 1 (33) |
| Class_HP | N | N | Y |
| AD | D | H | D |
| Years | 5.7167 | 3.963 | 7.1515 |
| Age | 50.8167 | 50.1111 | 51.3939 |
| Region | U | U | U |
| Gender | M | M | F |
| height | 160.8 | 167.4444 | 155.3636 |
| weight | 65.9333 | 69.4444 | 63.0606 |
| stomach_pain | N | N | N |
| F_Suffered | Y | N | Y |
| F_Expired | N | N | Y |
| F_E_AD | N | N | N |
| Tea | Y | Y | Y |
| Cofee | N | N | N |
| milk | N | N | N |
| curd | Y | Y | Y |
| Smoking | N | N | N |
| Tobacco | N | N | N |
| Alcohol | N | N | N |
| Tension | Y | N | Y |
| Stress | Y | N | Y |
| Exercise | N | Y | N |
| Travel | Mk | Mk | Mk |
| sleep_T | 2164.8333 | 2101.1111 | 2216.9697 |
| wokeup_T | 587.25 | 572.7778 | 599.0909 |

Figure 1. Analysis by Cluster Centroids

Figure 1 was shown the data Analysis using Cluster Centroids. The data was predicted with more chances for people in Urban regions with family history suffers gastric, expiry due to heart stoke with gastric, drink tea, with tension and stress.

```

Alternating decision tree:

: 0.23
| (1)stomach_pain = Y: -1.427
| (1)stomach_pain = N: 0.403
| | (2)F_Suffered = Y: -0.291
| | (2)F_Suffered = N: 0.697
| | | (3)AD = CD: -0.57
| | | (3)AD != CD: 0.876
| | (7)Age < 29.5: -0.561
| | (7)Age >= 29.5: 0.423
| | | (9)Exercise = Y: -0.109
| | | (9)Exercise = N: 0.596
| (4)wokeup_T < 665: -0.257
| | (5)height < 169.5: -0.248
| | | (6)Years < 1.5: 0.435
| | | (6)Years >= 1.5: -0.478
| | | | (8)height < 157.5: 0.171
| | | | (8)height >= 157.5: -0.862
| | | | (10)Region = U: -0.62
| | | | (10)Region = R: 0.356
| | (5)height >= 169.5: 0.838
| (4)wokeup_T >= 665: 0.709
Legend: -ve = Y, +ve = N
Tree size (total number of nodes): 31
Leaves (number of predictor nodes): 21

Time taken to build model: 0.02 seconds
    
```

Figure 2. AD Tree

Figure 2 shows the prediction of *H.pylori* using AD Tree. The Legend is positive for *H.pylori* (N) showing stomach pain, family suffered previously and Age >29.5.

```

Attribute Subset Evaluator (supervised, Class (nominal): 1 Class_HP):
  CFS Subset Evaluator
  Including locally predictive attributes

Selected attributes: 9,12,17 : 3
  stomach_pain
  F_E_AD
  Smoking
    
```

Figure 3. Prediction of association with Subset Evaluator

Figure 3 was shown the association of attributes with the Subset Evaluator. The method was shown that Stomach pain, Aging diseases and smoking are the attributes associated with *H.pylori*.

Table 1 was shown that the highest accuracy for the data was predicted with random Forest and K Star followed by J48, Bayes Net, Naive Bayes, LMT and Simple Cost.

| Table 1. Data Classification | |
|------------------------------|--------------------------------|
| Accuracy Data Representation | |
| Classifier | Correctly Classified Instances |
| BayesNet | 88% |

| | |
|---------------|--------|
| Naive Bayes | 83.33% |
| KStar | 100% |
| J48 | 96.66% |
| LMT | 83.33% |
| Random Forest | 100% |
| Simple Cost | 83.33% |

H.pylori is related to several ageing diseases and presently one of the importance microorganism that was focussed by several scientists for understanding about relationship of *H.pylori* with human health. The protein interaction analysis was predicted that SGCG protein is associated with several other aging disease proteins and immune system like PSEN1, APO, BAC E1, IL7, IL5R etc. These studies were predicted that *H.pylori* was the species associated with aging disease.

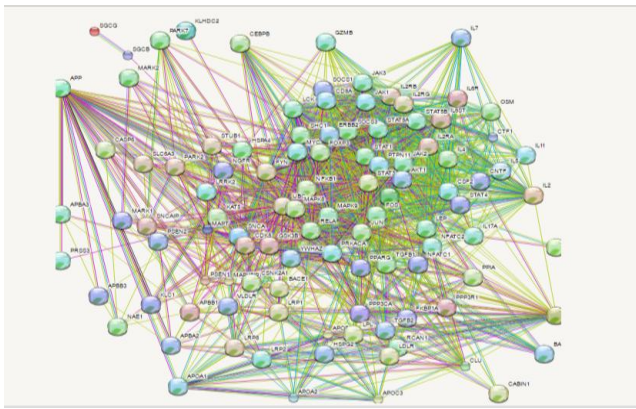


Figure 4. Protein Network Analysis

Helicobacter pylori cause a chronic bacterial infection associated with coronary heart disease and stroke in the older ages [13,14]. *H. pylori* mainly colonizes the gastric mucosa leads to infection is the main cause for the health problems like peptic ulcer and gastric malignancy [15]. *H. pylori*-associated diseases are the challenging tasks started during the latter half of the 20th century in epidemiology, treatment, and infection control of *H. pylori*.

Enhanced inflammation, accelerated cellular ageing or pathogen-dependent tissue destruction study mechanisms are important in the developed countries due to changes in metabolic and climatic changes. Intake of alcohol and antioxidant vitamins and smoking are associated with coronary heart disease [16].

V. CONCLUSION AND FUTURE SCOPE

The present work concludes that based on machine learning approaches there is a relationship of several ageing disease with *Helicobacter Pylori*. A network approach has been constructed for *H.pylori* proteins related proteins with aging disease. Further studies have to be carried on proteins network and their relation with system approaches.

ACKNOWLEDGMENT

The authors acknowledged to thank Management and Staff of University Teaching Departments (UTD) of Bilaspur University in bringing out the research work.

REFERENCES

- [1] N. M. Nasrabadi, "Pattern recognition and machine learning", Journal of electronic imaging, Vol. 16, Issue. 4, pp. 049901, 2007.
- [2] D. W. Aha, D. Kibler, M. K. Albert, "Instance-based learning algorithms", Machine learning, Vol.6, Issue. 1, pp. 37-66, 1991.
- [3] R. S. Michalski, "Learnable evolution model: Evolutionary processes guided by machine learning", Machine learning, Vol. 38, Issue. 1-2, pp. 9-40, 2000.
- [4] R. P. Kumar, M. Rao, D. Kaladhar, "Data categorization and noise analysis in mobile communication using machine learning algorithms", Wireless Sensor Network, Vol. 4, Issue. 4, pp. 113, 2012.
- [5] P. N. Rao, T. Devi, D. Kaladhar, G. Sridhar, A. A. Rao, "A Probabilistic Neural Network Approach For Protein Superfamily Classification", Journal of Theoretical & Applied Information Technology, Vol. 6, Issue. 1, pp.101-105, 2009.
- [6] A. Dubois, "Spiral bacteria in the human stomach: the gastric helicobacters", Emerging infectious diseases, Vol. 1, Issue. 3, pp. 79, 1995.
- [7] A. Mehmood, M. Akram, A. A. Shahab-uddin, K. Usmanhani, A. Hannan, E. Mohiuddin, M. Asif, "Helicobacter pylori: an introduction", International Journal of Applied Biology and Pharmaceutical Technology, Vol.1, Issue.3, pp. 1337-1351, 2010.
- [8] T. S. Bharath, M. S. Reddy, R. Dhanapal, N. G. R. Kumar, P. N. Raju, T. R. Saraswathi, "Molecular detection and correlation of Helicobacter pylori in dental plaque and gastric biopsies of dyspeptic patients", Journal of oral and maxillofacial pathology: JOMFP, Vol. 18, Issue. 1, pp. 19, 2014.
- [9] J. Lynch, G. D. Smith, "A life course approach to chronic disease epidemiology", Annu. Rev. Public Health, Vol. 26, pp. 1-35, 2005.
- [10] W. D. Chey, G. I. Leontiadis, C. W. Howden, S. F. Moss, "ACG clinical guideline: treatment of Helicobacter pylori infection", The American journal of gastroenterology, Vol. 112, Issue.2, pp. 212, 2017.
- [11] T. Farooqui, A. A. Farooqui, "Aging: an important factor for the pathogenesis of neurodegenerative diseases", Mechanisms of ageing and development, Vol.130, Issue.4, pp. 203-215, 2009.
- [12] J. Danesh, R. Collins, R. Peto, "Chronic infections and coronary heart disease: is there a link?", The lancet, Vol.350, Issue. 9075, pp. 430-436, 1997.
- [13] P. H. Whincup, M. A. Mendall, I. J. Perry, D. P. Strachan, M. Walker, "Prospective relations between Helicobacter pylori infection, coronary heart disease, and stroke in middle aged men", Heart, Vol.75, Issue.6, pp. 568-572, 1996.
- [14] G. Gavazzi, K. H. Krause, "Ageing and infection", The Lancet infectious diseases, Vol.2, Issue. 11, pp. 659-666, 2002.
- [15] D. Rothenbacher, H. Brenner, "Burden of Helicobacter pylori and H. pylori-related diseases in developed countries: recent developments and future implications", Microbes and Infection, Vol.5, Issue. 8, pp. 693-703, 2003.
- [16] M. G. Hertog, D. Kromhout, C. Aravanis, H. Blackburn, R. Buzina, F. Fidanza, Simona Giampaoli, Annemarie Jansen, Alessandro Menotti, Srecko Nedeljkovic, Maija Pekkarinen, B. S. Simic, Hironori Toshima, E. J. M. Feskens, P. C. H. Hollman, M. B. Katan, "Flavonoid intake and long-term risk of coronary heart

disease and cancer in the seven countries study”, Archives of internal medicine, Vol.155, Issue. 4, pp. 381-386, 1995.

Authors Profile

Dr. DSVGK Kaladhar is a teacher, administrator and researcher serving for the academic and research developments in India. He has 19 years of teaching and research experience and is presently working as Associate Professor, Dept. of Microbiology and Bioinformatics, Bilaspur University (a state Govt. University), Bilaspur (C.G.). He did M.Sc. in Microbiology, MCA, PhD in Biotechnology, PGDCA and Diploma in Bioinformatics. As an academician, Dr.Kaladhar doing duties like HOD Microbiology and Bioinformatics, Coordinator- PG Diploma in Yoga and PG Diploma in Computational Biology Programs etc. from Bilaspur University (CG). Dr. Kaladhar is also acting as Advisory Board Members, Academic Committee Member etc., in colleges under BU, SARC, etc. He was contributed more than 150 research papers including Elsevier, springer, nature, Bentham and many more. He was written and published about 15 books like Molecular Biochemistry, BioJava, Industrial Biotechnology etc. Dr.Kaladhar was also guided 5 Ph.D. scholars and more than 40 PG and UG students. Dr.Kaladhar is an Editor-in-chief to International Journal of Research Studies in Biosciences, Editorial Board member and Reviewer for several journals. Due to his contribution in teaching and research he received about 15 awards like *Bharat Jyoti Award*, *Best Citizen of India Award*, *Best Teacher and Researcher award*, *Bharat Siksha award*, *Raising son of India award*, etc.



Mrs Santosh Kumari Duppala pursued Bachelor of Science and Master of Science from Bilaspur University of Bilaspur (CG), India. She is currently pursuing Ph.D. and currently working as Chairman and Secretary in MRTI (an NGO), Bilaspur (CG), India. She has published more than 5 research papers in reputed international journals Her main research work focuses on Microbiology, Machine Learning, Biotechnology and Computational Biology. She has 3 years of Research Experience.