



CLASSIFICATION OF HELICOBACTER PYLORI ACCORDING TO NATIONAL STRAINS USING BAYESIAN LEARNING

Bekir Karlık, Alpaslan Avcı, A. Talha Yabanığül
Fatih University, Department of Computer Engineering, Istanbul, Turkey
bkarlik@fatih.edu.tr, alparslanavci@gmail.com, talhayabanigul@yahoo.com

Abstract- There is many studies about *Helicobacter pylori* genome and many instances of national strains are sequenced completely. To make a successful classification, the same functional portions have to be used a classifier like Bayesian Learning. Thus suitable genes will be used for classification since genes are portions that work functionally same. The *cagA* and *vacA* genes are selected for classification. *cagA* gene stands for ‘cytotoxin-associated protein A’ gene and *vacA* gene stands for ‘vacuolating cytotoxin precursor’ gene and these genes have a role of coding of these proteins. The reasons for selecting these genes are that these genes are the genes which affect the bacteria being a pathogen, Nucleotide numbers of these genes are higher than the most of other genes of bacteria, and these genes are the most popular genes of *Helicobacter pylori*. There are some instances of these genes which are classified with respect to national strains. The national strains are based on the nation of the host human. The cause of the difference on national strains is the difference of the cultural activities. If the national strain of a random *Helicobacter pylori* bacterium is known, the host human's nation can also be known approximately. The aim of this study is to classify *Helicobacter pylori* according to National strain using a well-know classification technique named Bayesian Learning.

Keywords- *Helicobacter pylori*, Bayesian Learning, National strain

1. INTRODUCTION

Helicobacter Pylori is a bacterium that infects the mucus lining of the stomach and duodenum. Many cases of peptic ulcers, gastritis, and duodenitis are caused by *H. pylori* infection. However, many who are infected do not show any symptoms of disease. *Helicobacter* bacteria are the only known microorganisms that can thrive in the highly acidic environment of the stomach. Its helical shape (hence the name *helicobacter*) is thought to have evolved to penetrate and colonize the mucus lining. The studies for sequencing the genome of *Helicobacter pylori* are completed and several strains are known as now. The genome of the strain "26695" consists of about 1.7 million base pairs, with some 1550 genes. The two sequenced strains show large genetic differences, with up to 6% of the nucleotides differing.

Last decades, there were some related works about *Helicobacter pylori* national strains. One of them is *Helicobacter pylori* in North and South America before Columbus. It is presented a molecular epidemiologic study, based on an analysis of *vacA*, *cagA* and *cag* right end junction genotypes from 1042 *Helicobacter pylori* isolates,

suggesting that *H. pylori* was present in the New World before Columbus. Eight Native Colombian and Alaskan strains possessed novel *vacA* and/or *cagA* gene structures and were more closely related to East Asian than to non-Asian *H. pylori*. Some Native Alaskan strains appear to have originated in Central Asia and to have arrived after strains found in South America suggesting that *H. pylori* crossed the Bering Strait from Asia to the New World at different times [1].

A collection of 20 strains of *Helicobacter pylori* from several regions of the world was studied to better understand the population genetic structure and diversity of this species. Sequences of fragments from seven housekeeping genes (*atpA*, *efp*, *mutY*, *ppa*, *trpC*, *urel*, *yphC*) and two virulence-associated genes (*cagA*, *vacA*) showed high levels of synonymous sequence variation (mean percentage K_s of 10-27%) and lower levels of non-synonymous variation (mean percentage K_a of 0.2-5.6%). Cluster analysis of pair wise differences between alleles revealed the existence of two weakly clonally groupings, which included half of the strains investigated. All six strains isolated from Japanese and coastal Chinese were assigned to the 'Asian' clonally grouping, probably reflecting descent from a distinct common ancestor. The clonally groupings were not totally uniform; recombination, as measured by the homoplasy test and compatibility matrices, was extremely common within all genes tested, except *cagA*. The fact that colonial descent could still be discerned despite such frequent recombination possibly reflects founder effects and geographical separation and/or selection for particular alleles of these genes [2].

Study of the *H. pylori* genome is centered on attempts to understand pathogenesis, the ability of this organism to cause disease. However, the purpose of this project is about classification of *Helicobacter pylori*. As it is stated above there are several strains of the genome and national strains are a class of these strains. The national strains are based on the nation of the host human. The cause of the difference on national strains is the difference of the cultural activities. If the national strain of the random *Helicobacter pylori* bacteria is known, the host human's nation can also be known approximately. By using this technique a human whose nation is not known can be found.

The classification of DNA is difficult because the sequence of DNA differs from among all individuals. As a highly developed technique of classification Bayesian Learning can be used to solve this problem. So, this study presents a Naive Bayesian classifier to solve this problem successfully.

2. BACKGROUND

Helicobacter Pylori is a bacterium that infects the mucus lining of the stomach and duodenum. Many cases of peptic ulcers, gastritis, and duodenitis are caused by *H. pylori* infection. However, many who are infected do not show any symptoms of disease. *Helicobacter* bacteria are the only known microorganisms that can thrive in the highly acidic environment of the stomach. Its helical shape (hence the name *helicobacter*) is thought to have evolved to penetrate and colonize the mucus lining.

In 1875, German scientists found spiral bacteria in the lining of the human stomach; the bacteria could not be grown in culture and the results were eventually forgotten. In 1892, the Italian researcher Giulio Bizzozzero described spiral bacteria living in the acidic environment of the stomach of dogs.

Professor Walery Jaworski of the Jagiellonian University in Krakow investigated sediments of gastric washings obtained from humans in 1899. Among some rod-like bacteria, he also found bacteria with a characteristic spiral shape, which he called *Vibrio rugula*. He was the first to suggest a possible role of this organism in the pathogen of gastric diseases. This work was included in the "Handbook of Gastric Diseases" but it did not have much impact as it was written in Polish. The bacterium was rediscovered in 1979 by Australian pathologist Robin Warren, who did further research on it with Barry Marshall beginning in 1981; they isolated the organisms from mucosal specimens from human stomachs and were the first to successfully culture them. In their original paper, Warren and Marshall contended that most stomach ulcers and gastritis were caused by colonization with this bacterium, not by stress or spicy food as had been assumed before.

The medical community was slow to recognize the role of this bacterium in stomach ulcers and gastritis, believing that no bacterium could survive for long in the acidic environment of the stomach. The community began to come around after further studies were done, including one in which Marshall drank a Petri dish of *H. pylori*, developed gastritis, and the bacteria were recovered from his stomach lining, thereby satisfying three out of the four Koch's postulates. Marshall's gastritis later resolved without treatment. Marshall and Warren went on to show that antibiotics are effective in the treatment of gastritis [3].

In 1994, the National Institutes of Health (USA) published an opinion stating that most recurrent gastric ulcers were caused by *H. Pylori*, and recommended that antibiotics be included in the treatment regimen. Evidence has been accumulating to suggest that duodenal ulcers are also associated with *H. Pylori* infection. In 2005, Warren and Marshall were awarded the Nobel Prize in Medicine for their work on *H. pylori*. Before the appreciation of the bacterium's role, stomach ulcers were typically treated with medicines that neutralize stomach acid or decrease its production. While this worked well, the ulcers very often reappeared. A traditional medication against gastritis was bismuth subsalicylate. It was often effective, but fell out of use, since its mechanism of action was a mystery. Nowadays it is quite clear that it is due to the bismuth salt acting as an antibiotic. Today, many stomach ulcers are treated with antibiotics effective against *H. pylori*. The bacterium was initially named *Campylobacter pyloridis*, then *C. pylori* (after a correction to the Latin grammar) and in 1989, after DNA sequencing and other data showed that the bacterium did not belong in the *Campylobacter* genus, it was placed in its own genus, *Helicobacter*. The name *Pylori* comes from the Greek word *pylorus*, which means gatekeeper, and refers to the pyloric valve (the circular opening leading from the stomach into the duodenum). While *H. pylori* remains the most important known bacteria to inhabit the human stomach, several other species of the *Helicobacter* genus have now been identified in other mammals and some birds, and some of these can infect humans. *Helicobacter* species have also been found to infect the livers of certain mammals and to cause liver disease [4].

H. Pylori are a spiral-shaped gram-negative bacterium, about 3 micrometers long with a diameter of about 0.5 micrometer. It has 4-6 flagella. It is microaerophilic, i.e. it requires oxygen but at lower levels than those contained in the atmosphere. It contains a hydrogenase and obtains energy by oxidizing molecular hydrogen (H_2) that was produced by other intestinal bacteria. It tests positive for oxidase and catalase.

Molecular model of *H. Pylori* urea's enzyme with its flagella and its helical shape, the bacterium drills into the mucus layer of the stomach, and then can be found in a number of locations: in the mucus, attached to epithelial cells, or inside vacuoles in epithelial cells. It produces adhesions which bind to membrane-associated lipids and carbohydrates and help its adhesion to epithelial cells. It excretes the enzyme urea's, which converts urea into ammonia and bicarbonate. The release of ammonia is beneficial to the bacterium since it partially neutralizes the very acidic environment of the stomach (whose very purpose is to kill bacteria). Ammonia is, however, toxic to the epithelial cells, and with other products of *H. pylori*, including protease, catalyses, and phospholipases, cause damage to those cells. Some strains of the bacteria have a particular mechanism for "injecting" the inflammatory agent peptidoglycan from their own cell wall into epithelial stomach cells. It remains unknown how this mechanism is advantageous to the bacterium. Under conditions of environmental stress, *Helicobacter* will convert from the spiral to a coccid form. This coccid form of the organism has not been cultured, but has been found in the water supply in the US and is apparently involved in the epidemiology of the bacterium. The coccid form has also been found to be able to adhere to gastric epithelial cells in vitro.

3. INFECTION AND DIAGNOSIS AND TREATMENT

Infection may be symptomatic or asymptomatic (without visible ill effects). It is estimated that up to 70% of infection is asymptomatic. The bacteria have been isolated from feces, saliva and dental plaque of infected patients, which suggests gastro-oral or fecal-oral as possible transmission routes. It is estimated that about 2/3 of the world population are infected by the bacterium. Actual infection rates vary from nation to nation - the West (Western Europe, North America, and Australasia) having rates around 25% and the Third World much higher. In the latter, it is common, probably due to poor sanitary conditions, to find infections in children. In the United States, infection is primarily in the older generations (about 50% for those over the age of 60 compared with 20% under 40 years) and the poorest. This is largely attributed to higher hygiene standards and widespread use of antibiotics. However, antibiotic resistance is appearing in *H. pylori*. There are already many metronidazole resistant strains in Europe, the United States, and developing countries. It is widely believed that in the absence of treatment, *H. pylori* infection persists for life; the human immune system is not able to eradicate it. However, despite the dominance of this belief among physicians, there is actually no epidemiological evidence to support it and increasing evidence to the contrary. Because *H. pylori* infection is not generally detected at onset or during the acute phase, the proportion of acute infections that persist is not known, but several studies that followed the natural history in populations have reported apparent spontaneous elimination. One can test for *H. pylori* infection with blood antibody or stool antigen tests, or with the carbon urea breath test (in which the patient drinks ¹⁴C- or ¹³C-labelled urea, which the bacterium metabolizes producing labeled carbon dioxide that can be detected in the breath), or endoscope to provide a biopsy sample for testing for the presence of urea's "rapid urea's test", histology or microbial culture. None of these test methods are completely failsafe. Blood antibody tests, for example, range from 76% to 84%

sensitivity. Medication can affect H. pylori urea's activity and give "false negatives" with the urea-based tests.

In patients who are asymptomatic, treatment is not usually recommended. In gastric ulcer patients where H. pylori are detected, normal procedure is eradication to allow the ulcer to heal. The standard first-line therapy is a one week triple-therapy. The Sydney gastroenterologist Thomas Borody invented the first triple therapy in 1987. Today the standard triple therapy is amoxicillin; clarithromycin and a proton pump inhibitor such as omeprazole -though sometimes a different proton pump inhibitor is substituted, or metronidazole is used in place of amoxicillin in those allergic to penicillin. Such a therapy has revolutionized the treatment of gastric ulcers and has made a cure to the disease possible, where previously symptom-control using antacids, H₂-antagonists or proton pump inhibitors alone was the only option. Unfortunately, an increasing number of infected individuals are found to harbor bacteria resistant to first-line antibiotics. This results in initial treatment failure and requires additional rounds of antibiotic therapy. For resistant cases, a quadruple therapy may be used. Bismuth compounds are also effective in combination with the above drugs. For the treatment of clarithromycin-resistant strains of H. pylori the use of levofloxacin as part of the therapy has been recommended.

There is some preliminary evidence that regular consumption of broccoli sprouts might eradicate H. pylori. Some evidence suggested that consumption of mastic gum might be able to control or even eradicate H. pylori, but later studies showed this not to be the case. A study done on Mongolian gerbils indicates that green tea extract can suppress H. pylori growth. Another study done in South Korea suggests that the acidic polysaccharide found in green tea was significantly effective in preventing adhesion of H. pylori cells to human cultures of epithelial cells [5].

4. GENOME STUDIES OF DIFFERENT STRAINS

Several strains are known, and the genomes of two have been completely sequenced. The genome of the strain "26695" consists of about 1.7 million base pairs, with some 1550 genes. The two sequenced strains show large genetic differences, with up to 6% of the nucleotides differing. Study of the H. pylori genome is centered on attempts to understand pathogenesis, the ability of this organism to cause disease. There are 62 genes in the "pathogenesis" category of the genome database. Both sequenced strains have an approximately 40 kb long Cag pathogenicity island (a common gene sequence believed responsible for pathogenesis) that contains over 40 genes. This pathogenicity island is usually absent from H. pylori strains isolated from humans who are carriers of H. pylori but remain asymptomatic.

The *cagA* gene codes for one of the major H. pylori virulence proteins. Bacterial strains that have the *cagA* gene are associated with an ability to cause severe ulcers. The *cagA* gene codes for a relatively long (1186 amino acid) protein. The CagA protein is transported into human cells where it may disrupt the normal functioning of the cytoskeleton. The Cag pathogenicity island has about 30 genes that code for a complex type IV secretion system. After attachment of H. pylori to stomach epithelial cells, the CagA protein is injected into the epithelial cells by the type IV secretion system. The CagA protein is phosphorylated on tyrosine residues by a host cell membrane-associated tyrosine kinase. Pathogenic strains of H. pylori have been shown to activate the

epidermal growth factor receptor (EGFR), a membrane protein with a tyrosine kinase domain. Activation of the EGFR by *H. pylori* is associated with altered signal transduction and gene expression in host epithelial cells that may contribute to pathogenesis. It has also been suggested that a c-terminal region of the *cagA* protein (amino acids 873-1002) can regulate host cell gene transcription independent of protein tyrosine phosphorylation. It is thought, due to *cagA*'s low GC content relative to the rest of the helicobacter genome, that the gene was acquired by horizontal transfer from another *cagA*+ bacterial species.

Each human population has a characteristic distribution of *H. pylori* strains that typically infect members of that population. This allows researchers to use *H. pylori* to study human migration patterns. It could be established that *H. pylori* in Amazon Indians has East Asian rather than European origins, suggesting that it arrived with the original immigrants at least 11,000 years ago [1, 5].

5. NAIVE BAYES CLASSIFIER

A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be independent feature model. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without believing in Bayesian probability or using any Bayesian methods [6-9].

In spite of their naive design and apparently over-simplified assumptions, naive Bayes classifiers often work much better in many complex real-world situations than one might expect. Recently, careful analysis of the Bayesian classification problem has shown that there are some theoretical reasons for the apparently unreasonable efficacy of naive Bayes classifiers. An advantage of the Naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix. Abstractly, the probability model for a classifier is a conditional model

$$p(C|F_1, \dots, F_n) \quad (1)$$

over a dependent class variable C with a small number of outcomes or classes, conditional on several feature variables F_1 through F_n . The problem is that if the number of features n is large or when a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it more tractable. Using Bayes' theorem, we can write;

$$p(C|F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)} \quad (2)$$

In plain English the above equation can be written as;

$$\text{Posterior} = \text{Prior} * \text{Likelihood} / \text{Evidence} \quad (3)$$

In practice, we are only interested in the numerator of fraction, since the denominator does not depend on C and the values of the features F_i are given, so that the

denominator is effectively constant. The numerator is equivalent to the joint probability model

$$p(C, F_1, \dots, F_n) \quad (4)$$

this can be rewritten as follows, using repeated applications of the definition of conditional probability and so forth. Now the "naive" conditional independence assumptions come into play: assume that each feature F_i is conditionally independent of every other feature F_j for $j \neq i$. This means that

$$p(F_i|C, F_j) = p(F_i|C) \quad (6)$$

and so the joint model can be expressed as;

$$p(C, F_1, \dots, F_n) = p(C)p(F_1|C)p(F_2|C)p(F_3|C)\dots = p(C)\prod_{i=1}^n p(F_i|C) \quad (7)$$

This means that under the above independence assumptions, the conditional distribution over the class variable C can be expressed like this:

$$p(C|F_1, \dots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^n p(F_i|C) \quad (8)$$

where Z is a scaling factor as constant, if the values of the feature variables are known. Models of this form are much more manageable, since they factor into a so-called class prior $p(C)$ and independent probability distributions. If there are k classes and if a model for $p(F_i)$ can be expressed in terms of r parameters, then the corresponding naive Bayes model has $(k-1)+nr$ parameters. In practice, often $k=2$ (binary classification) and $r=1$ (Bernoulli variables as features) are common, and so the total number of parameters of the naive Bayes model is $2n+1$, where n is the number of binary features used for prediction [10].

6. MATERIAL AND METHODS

To achieve the goals of the study has to be provided a number of functional requirements. There are many studies about Helicobacter pylori genome and many instances of national strains are sequenced completely. Thus suitable genes will be used for classification since genes are portions that work functionally same.

The *cagA* and *vacA* genes are selected for classification. *cagA* gene stands for "cytotoxin-associated protein A" gene and *vacA* gene stands for "vacuolating cytotoxin precursor" gene and these genes have a role of coding of these proteins. The reasons for selecting these genes are that:

- These genes are the genes which affect the bacteria being a pathogen
- Nucleotide numbers of these genes are higher than the most of the other genes of bacteria
- These genes are the most popular genes of Helicobacter pylori

There are some instances of these genes which are classified with respect to national strains. One of the studies about this classification is Helicobacter pylori in North and South America before Columbus [4]. This study will be used as taking the names of the instances of classified strains. PubMed Central is a free digital archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health (NIH) developed and managed by NIH's National Center for Biotechnology Information (NCBI) in the National Library of Medicine (NLM). With PubMed Central, NLM is

taking the lead in preserving and maintaining unrestricted access to the electronic literature, just as it has done for decades with the printed biomedical literature. PubMed Central aims to fill the role of a world class library in the digital age. It is not a journal publisher. NLM believes that giving all users free and unrestricted access to the material in PubMed Central is the best way to ensure the durability and utility of the archive as technology changes over time. Entrez is the integrated, text-based search and retrieval system used at NCBI for the major databases, including PubMed, Nucleotide and Protein Sequences, Protein Structures, Complete Genomes, Taxonomy, and others. In this study, classification data was taken from Entrez [4]. Here, the codes of the nucleotides in the genes are used letters of alphabet such as A, C, G, T which are not numerical values. To prevent this handicap the ASCII codes of these letters will be used as input values. The *cagA* and *vacA* genes have approximately 450 to 650 nucleotides in their sequence. All of these numbers of nucleotide is used as inputs of Bayesian classifier and all of these codes have to be converted to numerical values. There are some kinds of computer programs about Bayes classification and RapidMiner is one of that programs. The usage of that program is very simple. RapidMiner can make classifications by using different methods. So RapidMiner is used for Bayes classification [11].

7. RESULTS AND CONCLUSIONS

RapidMiner (formerly YALE) is the most comprehensive open-source software for intelligent data analysis, data mining, knowledge discovery, machine learning, predictive analytics, forecasting, and analytics in business intelligence. RapidMiner is implemented in Java and available under GPL (GNU General Public License) as well as under a developer license (OEM license) for closed-source developers.

Firstly, we select the root on the program and than input-output operator is selected by using “New Operator” option. After that, Example Source is selected on the Examples panel. Example Source is selected here because; we put our training data to application by from that Example Source option (see Figure 1). Next step is applying the classification algorithm to that selected example source so; we select a learner for the program. This learner is the algorithm of the training data. On the “New Operator” panel, we select “Learner” and than we choose the learning type. For our classification we choose Supervised learning. Than we choose the Bayes option from the Weka panel and after that W-Naive Bayes is selected (see Fig. 2). As it can be seen on Fig. 3, we put another example source for the program. Another example source is necessary because, we give our test data for the program and this test data is given as input-output operation. For these example sets which we selected for the program, need a Model Applier to apply training data and test data. That Model Applier combines all of the data which we give to the program and it gives us the results about out test data. After selecting the Model Applier, program gives us the results about the test data. As it can be seen on Fig.4, RapidMiner gives us the expected results of test data.

As a conclusion, our project is a classification implementation of human nations with respect to helicobacter pylori which can be found in the stomach. The *vacA* gene and *cagA* gene are differed among human hosts from different nations. So by looking these genes we can classify a human nation. When the national strain of the random helicobacter pylori is known, the host human’s nation can also be known approximately.

Our project uses that technique to find the nation of a human whose nation is not known. In this study, we obtained very high classification rate approximately 99 %. This classification can be used for classifying any genes of human on future.

REFERENCES

1. Achtman et al., Recombination and clonal groupings within Helicobacter pylori from different geographical regions, PMID: 10320570, <http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db>
2. Ji-Hye Lee, Jin Sun Shim, Jung Sun Lee, Jin Kwang Kim, In Seok Yang,, Mi-Sook Chung, and Kyung Hyun Kim, “Inhibition of Pathogenic Bacterial Adhesion by Acidic Polysaccharide from Green Tea (Camellia sinensis), *Journal of Agric. Food Chem.*, **54** 23, 8717 – 8723, 2006.
3. Martin J. Blaser, An Endangered Species in the Stomach Is the decline of Helicobacter pylori, a bacterium living in the human stomach since time immemorial, good or bad for public health?, February 2005. <http://www.sciam.com/article.cfm?articleID=00024948-2764-11E8-A28583414B7F00>
4. T.J. Borody, P. Cole, S. Noonan, A. Morgan, J. Lenne, L. Hyland, S. Brandl, E. G. Borody, L.L. George, Recurrence of duodenal ulcer and Campylobacter pylori infection after eradication, PMID: 2687668, 1999. <http://www.ncbi.nlm.nih.gov/sites/entrezcmd=retrieve&db=pubmed&listuids=2687668>
5. J.W. Olson and R. J. Maier, Molecular Hydrogen as an energy source for Helicobacter pylori, *Science*, 298.5599, 1788–1790, 2002. http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12459589
6. R. Winkler, The assessment of prior distributions in Bayesian analysis, *American Statistical Association Journal*, 62: 776-800, 1967.
7. H. Suermondt, G. Cooper, A combination of exact algorithms for inference on Bayesian belief networks, *Inter. Journal of Approximate Reasoning*, 5, 521-542, 1991.
8. D. Spiegelhalter, et al, Bayesian analysis in expert systems, *Statistical Science*, 8, 219-282, 1993.
9. M. Singh, and G. Provan, Efficient learning of selective Bayesian network classifier, Technical Report, MS-CIS-95-36, Computer and Information Science Department, University of Pennsylvania, PA, 1995.
10. F. Jensen, *An Introduction to Bayesian Networks*, Springer, 1996.
11. <http://rapid-i.com/content/view/26/82/lang,en/>

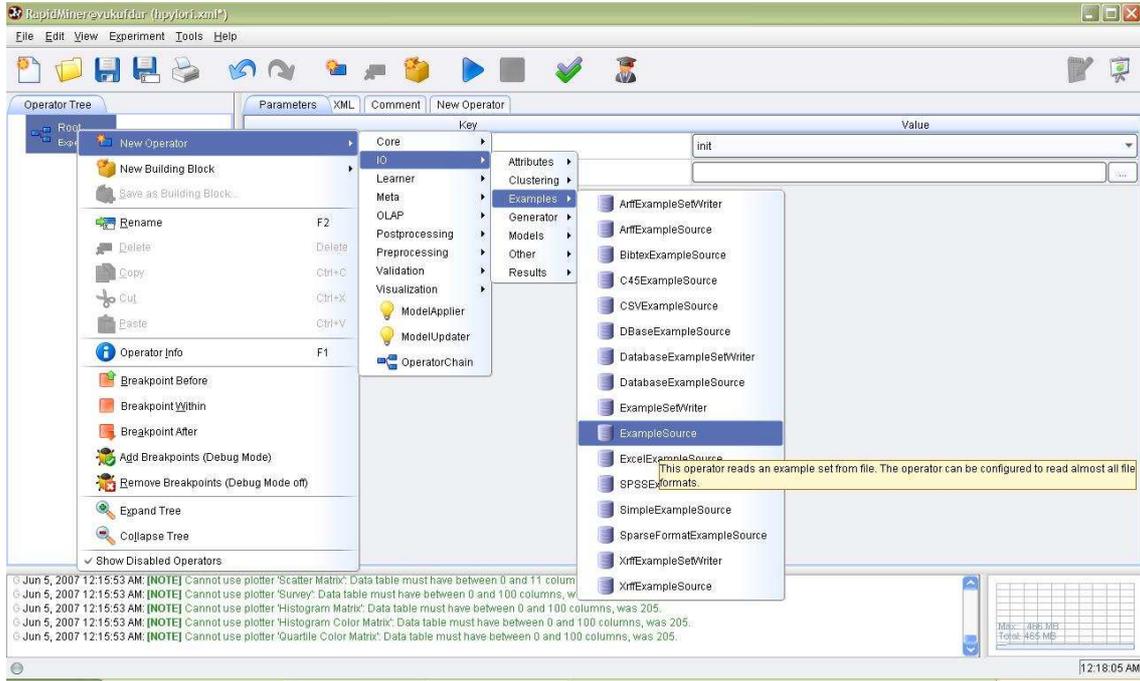


Fig. 1. Example source option

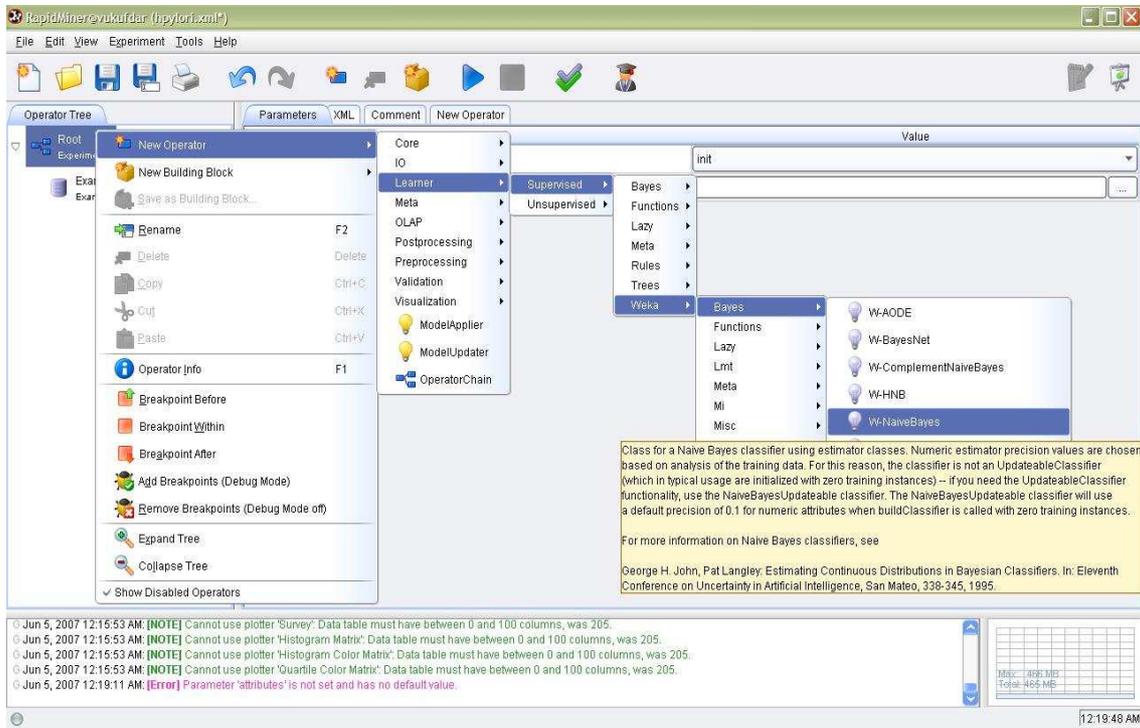


Fig. 2. Selection of W-Naive Bayes

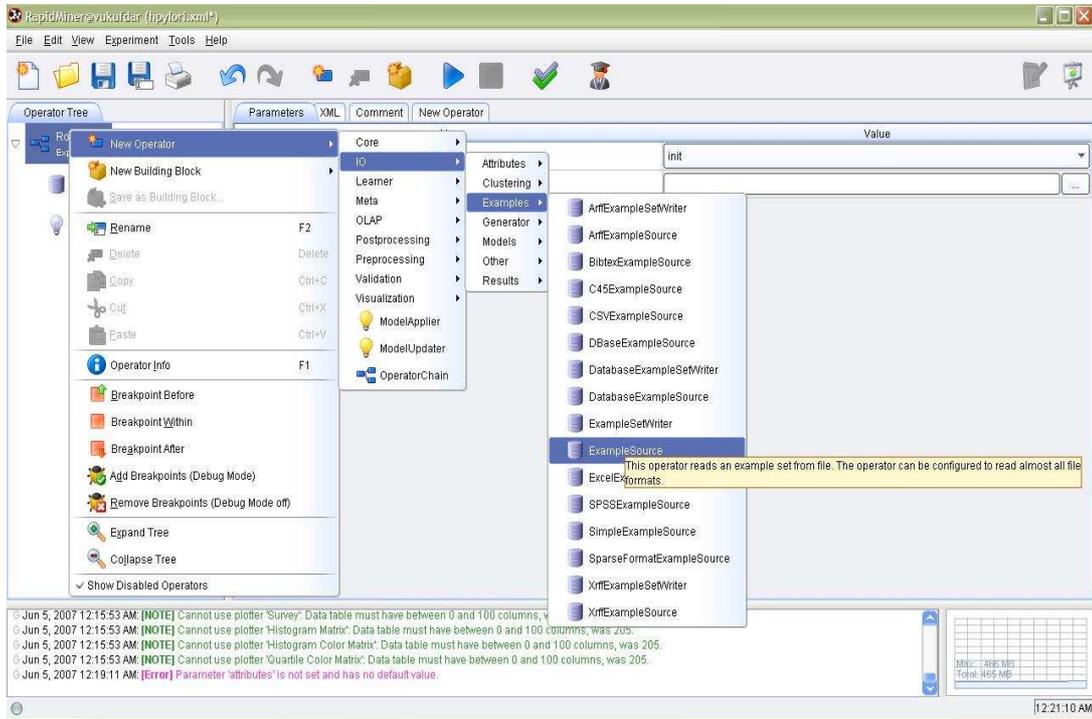


Fig. 3. Test phase

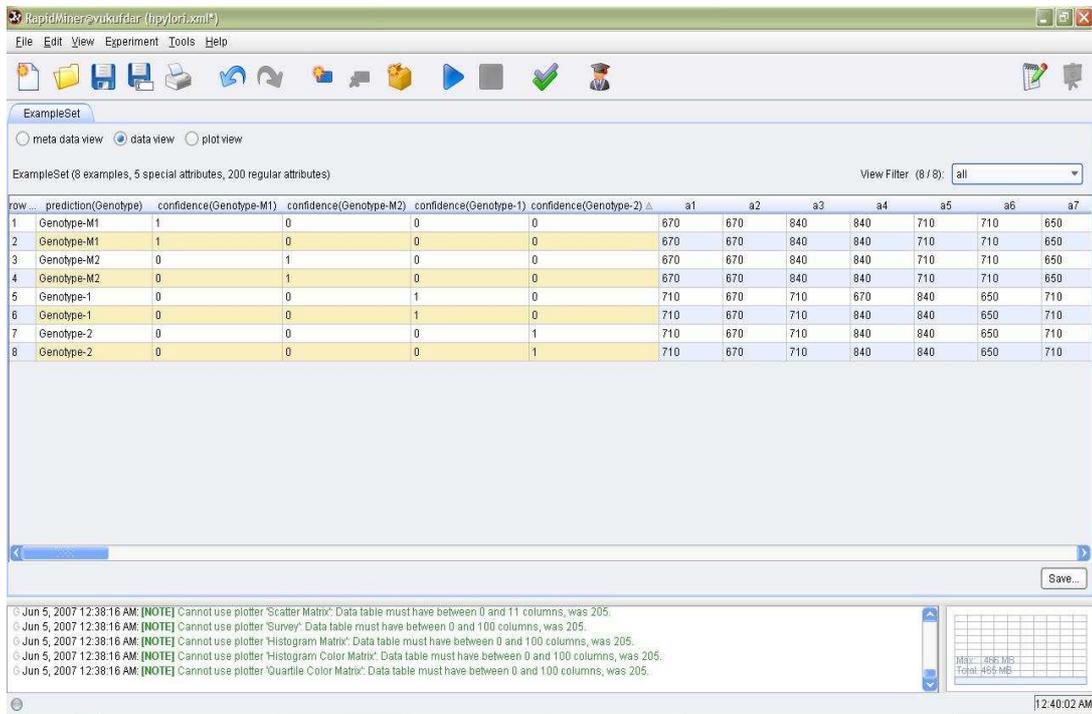


Fig. 4. The expected results of test data