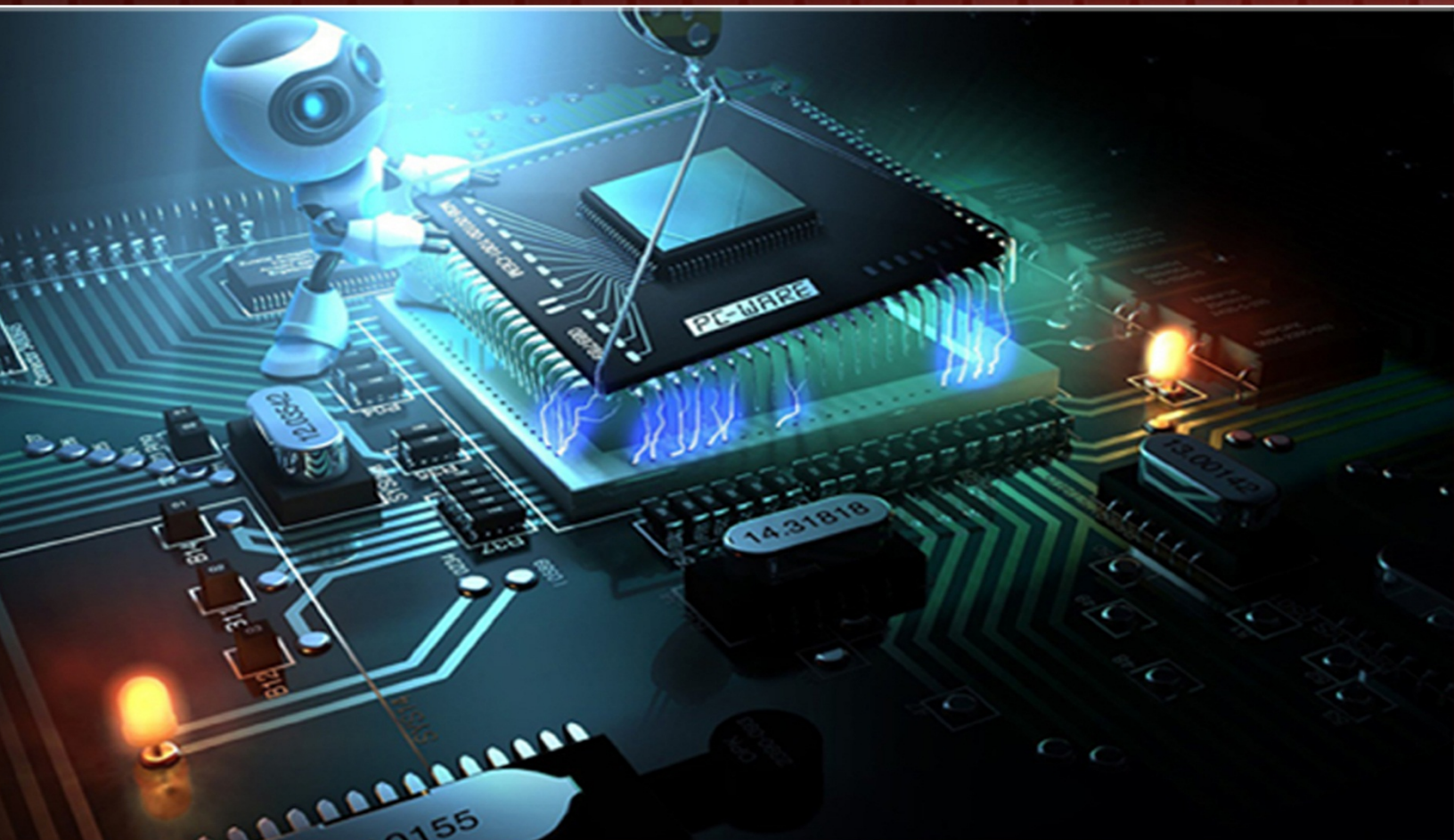


# TRANSACTIONS ON MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE



---

## TABLE OF CONTENTS

EDITORIAL ADVISORY BOARD	I
DISCLAIMER	II
<b>On Comparative Study for Two Selective Ant Colony's Optimal Decisions Versus Reconstruction Problem Solving by a Mouse Inside Figure of Eight (8) Maze</b> Hassan M. H. Mustafa, Fadhel Ben Tourkia	1
<b>Application of an Artificial Neural Network for Early and Accurate Diagnosis of Parkinson's disease</b> Akshar Agarwal	19
<b>An Ocean of Opportunities in Artificial Human Optimization Field</b> Satish Gajawada	24
<b>Determinants of Environmental Sanitation in the City of Tshwane, South Africa</b> Yohannes Worku and Mammo Muchie	32
<b>Classification of Encrypted Texts using Deep Learning</b> Zeinab Nazemi Absardi, Reza Javidan	49

---

## EDITORIAL ADVISORY BOARD

### Editor-in-Chief

**Professor Er Meng Joo**  
Nanyang Technological University  
*Singapore*

### Members

**Professor Djamel Bouchaffra**  
Grambling State University, Louisiana  
*United States*

**Prof Bhavani Thuraisingham**  
The University of Texas at Dallas  
*United States*

**Professor Dong-Hee Shin,**  
Sungkyunkwan University, Seoul  
*Republic of Korea*

**Professor Filippo Neri,**  
Faculty of Information & Communication Technology  
University of Malta  
*Malta*

**Prof Mohamed A Zohdy,**  
Department of Electrical and Computer Engineering  
Oakland University  
*United States*

**Dr Kyriakos G Vamvoudakis,**  
Dept of Electrical and Computer Engineering  
University of California Santa Barbara  
*United States*

**Dr Luis Rodolfo Garcia**  
College of Science and Engineering  
Texas A&M University, Corpus Christi  
*United States*

**Dr Hafiz M. R. Khan**  
Department of Biostatistics  
Florida International University  
*United States*

**Dr. Xuewen Lu**  
Dept. of Mathematics and Statistics  
University of Calgary  
*Canada*

**Dr. Marc Kachelriess**  
X-Ray Imaging and Computed Tomography  
German Cancer Research Center  
*Germany*

**Dr. Nadia Pisanti**  
Department of Computer Science  
University of Pisa  
*Italy*

**Dr. Frederik J. Beekman**  
Radiation Science & Technology  
Delft University of Technology, *Netherlands*

**Professor Wee SER**  
Nanyang Technological University  
*Singapore*

**Dr Xiaocong Fan**  
The Pennsylvania State University  
*United States*

**Dr Julia Johnson**  
Dept. of Mathematics & Computer Science  
Laurentian University, Ontario  
*Canada*

**Dr Chen Yanover**  
Machine Learning for Healthcare and Life Sciences  
IBM Haifa Research Lab  
*Israel*

**Dr Vandana Janeja**  
University of Maryland, Baltimore  
*United States*

**Dr Nikolaos Georgantas**  
Senior Research Scientist at INRIA, Paris-Rocquencourt  
*France*

**Dr Zeyad Al-Zhour**  
College of Engineering, The University of Dammam  
*Saudi Arabia*

**Dr Zdenek Zdrahal**  
Knowledge Media Institute,  
The Open University, Milton Keynes  
*United Kingdom*

**Dr Farouk Yalaoui**  
Institut Charles Dalaunay  
University of Technology of Troyes  
*France*

**Dr Jai N Singh**  
Barry University, Miami Shores, Florida  
*United States*

**Dr. Laurence Devillers**  
Computer Science, Paris-Sorbonne University  
*France*

**Dr. Hans-Theo Meinholz**  
Systems analysis and middleware  
Fulda University of Applied Sciences  
*Germany*

**Dr. Katsuhiko Honda**  
Department of Computer science and Intelligent Systems  
Osaka Prefecture University  
*Japan*

**Dr. Uzay Kaymak**  
Department of Industrial Engineering & Innovation  
Sciences, Technische Universiteit Eindhoven University of  
Technology, *Netherlands*

- 
- Dr. Ian Mitchell**  
Department of Computer Science  
Middlesex University London  
*United Kingdom*
- Dr. Weiru Liu**  
Department of Computer Science  
University of Bristol  
*United Kingdom*
- Dr. Aladdin Ayesh**  
School of Computer Science and Informatics  
De Montfort University, Leicester  
*United Kingdom*
- Dr. David Glass**  
School of Computing, Ulster University  
*United Kingdom*
- Dr. Sushmita Mukherjee**  
Department of Biochemistry  
Weil Corner Medical College, New York  
*United States*
- Dr. Rattikorn Hewett**  
Dept. of Computer Science  
Texas Tech University  
*United States*
- Dr. Cathy Bodine**  
Department of Bioengineering  
University of Colorado  
*United States*
- Dr. Daniel C. Moos**  
Education Department  
Gustavus Adolphus College  
*United States*
- Dr. Anne Clough**  
Department of Mathematics,  
Statistics and Computer Science, Marquette University  
*United States*
- Dr. Jay Rubinstein**  
University of Washington  
*United States*
- Dr. Frederic Maire**  
Department of Electrical Engineering and Computer Science  
Queensland University of Technology  
Australia
- Dr. Bradley Alexander**  
School of Computer Science  
University of Adelaide  
*Australia*
- Dr. Erich Peter Klement**  
Department of Knowledge-Based Mathematical Systems  
Johannes Kepler University Linz  
*Austria*
- Dr. Ibrahim Ozkan**  
Department of Economics
- Dr. Fernando Beltran**  
University of Auckland Business School  
*New Zealand*
- Dr. Mikhail Bilenko**  
Machine Intelligence Research (MIR) Group, Yandex  
*Russia*
- Anne Hakansson**  
Department of Software and Computer systems  
KTH Royal Institute of Technology  
*Sweden*
- Dr. Adnan K. Shaout**  
Department of Electrical and Computer Engineering  
University of Michigan-Dearborn  
*United States*
- Dr. Tomasz G. Smolinski**  
Department of Computer and Information Sciences  
Delaware State University  
*United States*
- Dr. Yi Ming Zou**  
Department of Mathematical Sciences  
University of Wisconsin  
*United States*
- Mohamed A. Zohdy**  
Department of Electrical and Systems Engineering  
Oakland University  
*United States*
- Dr. Krysta M. Svore**  
Microsoft Quantum – Redmond Microsoft  
*United States*
- Dr. John Platt**  
Machine learning, Google  
*United States*
- Dr. Wen-tau Yih**  
Natural language processing  
Allen Institute for Artificial Intelligence  
*United States*
- Dr. Matthew Richardson**  
Natural Language Processing Group, Microsoft  
*United States*
- Amer Dawoud**  
Department of Computer Engineering  
University of Southern Mississippi  
*United States*
- Dr. Jinsuk Baek**  
Department of Computer Science  
Winston-Salem State University  
*United States*
- Dr. Harry Wechsler**  
Department of Computer Science  
George Mason University  
*United States*



---

Hacettepe University  
*Canada*

**Dr. Sattar B. Sadkhan**  
Department of Information Networks  
University of Babylon  
*Iraq*

**Dr. Marina Papatriantafilou**  
Department of Computer Science and Engineering  
Chalmers University of Technology  
*Sweden*

**Dr. Florin Manea**  
Dependable Systems Group, Dept. of Computer Science  
Kiel University Christian-Albrechts  
*Germany*

**Prof. Dr. Hans Kellerer**  
Department of Statistics and Operations Research  
University of Graz  
*Austria*

**Dr. Dimitris Fotakis**  
School of Electrical and Computer Engineering  
National Technical University of Athens  
*Greece*

**Dr. Faisal N. Abu-Khzam**  
Department of Computer Science and Mathematics  
Lebanese American University, Beirut  
*Lebanon*

**Dr. Tatsuya Akutsu**  
Bioinformatics Center, Institute for Chemical Research  
Kyoto University, Gokasho  
*Japan*

**Dr. Francesco Bergadano**  
Dipartimento di Informatica,  
Università degli Studi di Torino  
*Italy*

**Dr. Mauro Castelli**  
NOVA Information Management School (NOVA IMS),  
Universidade Nova de Lisboa  
*Portugal*

**Dr. Stephan Chalup**  
School of Electrical Engineering and Computing,  
The University of Newcastle  
*Australia*

**Dr. Louxin Zhang**  
Department of Mathematics, National University of  
*Singapore*

**Dr. Omer Weissbrod**  
Department of Computer Science  
Israel Institute of Technology  
*Israel*

**Dr. Yael Dubinsky**  
Department of Computer Science  
Israel Institute of Technology  
*Israel*

**Dr. Francesco Bergadano**  
Department of Computer Science  
University of Turin  
*Italy*

**Dr. Marco Chiarandini**  
Department of Mathematics and Computer Science,  
University of Southern  
*Denmark*

**Dr. Xiaowen Chu**  
Department of Computer Science, Hong Kong Baptist  
University, Kowloon Tong  
*Hong Kong*

**Dr. Alicia Cordero**  
Instituto de Matemática Multidisciplinar, Universitat  
Politècnica de València  
*Spain*

**Dr. Sergio Rajsbaum**  
Instituto de Matemáticas, Universidad Nacional  
Autónoma de México  
*Mexico*

**Dr. Tadao Takaoka**  
College of Engineering  
University of Canterbury, Christchurch  
*New Zealand*

**Dr. Bruce Watson**  
FASTAR Group, Information Science Department,  
Stellenbosch University  
*South Africa*

**Dr. Tin-Chih Toly Chen**  
Department of Industrial Engineering and Management,  
National Chiao Tung University, Hsinchu City  
*Taiwan*

---

## **DISCLAIMER**

All the contributions are published in good faith and intentions to promote and encourage research activities around the globe. The contributions are property of their respective authors/owners and the journal is not responsible for any content that hurts someone's views or feelings etc.

# On Comparative Study for Two Selective Ant Colony's Optimal Decisions Versus Reconstruction Problem Solving by a Mouse Inside Figure of Eight (8) Maze

Hassan M. H. Mustafa, Fadhel Ben Tourkia

*Computer Engineering Department, Al-Baha Private College of Sciences Al-Baha,  
Kingdom of Saudi Arabia*

[prof.dr.hassanmoustafa@gmail.com](mailto:prof.dr.hassanmoustafa@gmail.com); [bintorkih@yahoo.com](mailto:bintorkih@yahoo.com)

## ABSTRACT

This piece of research addresses an interdisciplinary comparative study of two environmental challenging phenomenal issues. Both were associated to two nonhuman creatures characterized by their behavioral intelligent performance concerned with optimal diverse decisions paradigms. More specifically, this paper deals with the comparison study for analogical behavioral learning of social insects (Ants) colony performance, versus performed behavioral learning achievement by an animal considering a mouse inside Figure of eight (8) maze via its brain hippocampus "time cells" neurons.

In more precise details, this paper firstly have demonstrated for Ant Colony System (ACS) the two effective optimal selectivity decisions for : a) The best source location between two food sources that are equidistantly sited away from the original home nest, based upon pheromone trails and following the tandem running regulation & b) The balanced selection performance with the migration speed, in order to minimize exposure to a hostile environment to avoid vulnerability to presumable danger. Secondly, the optimal decisional issue is demonstrated related to mouse's behavioral learning intelligent approach which observed in practice following its active sequential trials aiming to reach the optimal solution for a reconstruction problem during its movement inside figure of eight (8) maze. Finally, after running of realistic Artificial Neural Networks' (ANNS) simulation programs. Interestingly, the obtained results considering models of both of suggested intelligent behavioral learning issues characterized by relevant functional analogy considering changed number of artificial ant agents versus the various number of neurons inside mouse's hippocampus brain area.

**Keywords-** Artificial neural network modeling, Swarm Intelligence, Tandem Running, House-HuntingAnts, Nature Inspired Computing.

## 1 Introduction

The Ant is one type of social insects that have been evolved from wasp-like ancestors in the mid-Cretaceous of the period between 110 and 130 million years ago [1]. It means they are as old as dinosaurs but unlike them, ants managed to survive. In more general sense, considering ants as social insects means they live in large colony systems or groups. In natural world, it is observable depending on the species, that ant colonies could be composed of millions ant members of (agents). [2] Accordingly, the colonies

of ants (as a social insects), are behaving adaptively while living in a competitive, and dynamical environment. Which characterized by constantly changing food sources in their location (distributed sites), and variation of their quantity and quality. Most of ant species are dependent upon ephemeral food finds. In such an environment, there is an advantage to sharing information if it can help the colony direct its workers quickly to the best food sources. The second paradigm considers collective decisional intelligence as a behavior that emerges through the interaction and cooperation of large numbers of lesser intelligent agents (such as ants). This paradigm composed of two dominant sub-fields 1) Ant Colony Optimization that investigates probabilistic algorithms inspired by the foraging behavior of ants [3][4][5], and 2) Particle Swarm Optimization that investigates probabilistic algorithms inspired by the flocking and foraging behavior of birds and fish [6]. Like evolutionary computation, swarm intelligence-based techniques are considered adaptive strategies and are typically applied to search and optimization domains. That simulation the foraging behavioral intelligence of a swarm (ant) system used for reaching optimal solution of TSP a cooperative learning approach to the traveling salesman problem optimal solution of TSP considered using realistic simulation of Non-neural systems namely: Ant Colony System ACS. In the context of intercommunications and cooperative learning among ants inside ACS. More specifically, Social insects have evolved impressively sophisticated solutions to some challenging environmental issues such as making nest site selection via building a leading model system of the collective intelligence of animal insect groups Temnothorax ants. Recently, educationalists have adopted interesting interdisciplinary challenging research work concerned with realistic modeling behavioral learning of some behavioral nonhuman creatures compared with ACS intelligent performance [7]. That's seems to be analogous to the natural aspects of behavioral ecological learning (swarm intelligence), observed by one type of social insects colonies namely (ACS) [6]. By referring to [6], therein has been interestingly stated that: " How do ants communicate so quickly and clearly? The German ethologist and socio biologist Prof. Bert Hölldobler researches in this field for many years. Together with the American Edward O. Wilson, he has revolutionized the study of ants. He says: An ant colony functions like a super organism. The life of each single being counts for little, the survival of the super organism, the state, is all what counts". Moreover, it has been announced, at[6] some of interesting findings associated to "Ants - Beautiful Creation of Allah", by following words "This particular insect has been selected as a title (Aayat) of a Surah No.27 "The Ants" in the Noble Qur'an. Briefly, it discusses the advent of Hadhrat Sulayman (Alayhis salaam), with his hosts of Jinn, men and birds and they marched in arranged groups till they reached the Valley of Ants. Said one Ant "Oh you Ants! Get back into your dwellings, lest Sulayman and his hosts crush you unintentionally". The Hadith and sayings of the companions are numerous relating to the Ant. Hadhrat Ibn Abbas (Radhiallaahu Anhu) reports that the Prophet (Pbuh) said, "Do not kill four of these creatures viz. ant, bee, sparrows (wood-pecker), and any pigeon". Hadhrat Abdur Rahman Ibn Abdullah (Radhiallaahu Anhu) reports that The Prophet (Pbuh) descended at a certain place - (house) and attended to the call of nature. Incidentally a man kindled a fire in a heap of ants, (either the house was built on a tree or on the ground). The Prophet (Pbuh) seeing this, exclaimed "Who did this? Extinguish it!! Extinguish it!" Abu Hurayrah (Radhiallaahu Anhu) reports that a certain Prophet from the many Prophets (Alayhis salaam) rested under the shade of a tree. An ant bit him. So the Prophet stood up to avenge for the bite. He spotted a large group of ants and killed them. A call was voiced to Him, "Was it not only one ant that bit you?" Abu Hurayrah (Radhiallaahu Anhu) reports, I heard the Prophet (Pbuh) saying, A certain Prophet from the many Prophets (Alayhis salaam) was resting under the shade of a tree. An ant bit him.



He instructed his belongings to be removed from there and immediately ordered the heap of Ants to be burnt. Allah Ta'ala revealed to him "Was it not only Ant that hurt you and you have ordered a full nation of ants to be destroyed who are engaged in the invocation of Allah Taala? Was it not one ant?" Hassam Dastawani says that the ants and small ants during the season of summer bring love and prosperity to us but during the season of winter we fear they will dissipate our hoards. Abu Bakr Najee reports that Hazrat Sulayman (Alayhis salaam) came out seeking water. Incidentally he spotted an ant with its bottom downwards and struggling with its legs up in the air upwards and saying, "Oh Allah we are one of your own creation, we are not independent of our thirst. Either you refreshen us and feed us or either kill us and destroy us." When they reached the Valley of the Ants, an Ant exclaimed, 'O Ants! Enter your dwellings lest Sulaymaan and his armies crush you without noticing it.' [Qurán 27:18] Life Style of Ants In recent times, research has shown us several facts about the lifestyle of ants, which were not known earlier to mankind. Research has shown that the animals or insects whose lifestyle is closest in resemblance of the lifestyle of human beings are the ants. This can be seen from the following findings regarding ants:

- a. The ants bury their dead in a manner similar to the humans.
- b. They have a sophisticated system of division of labour.
- c. Once in a while they meet among themselves to have a 'chat'.
- d. They have an advanced method of communication among themselves.
- e. They hold regular markets wherein they exchange goods.
- f. They store grains for long periods in winter and if the grain begins to bud, they cut the roots, If the grains stored by them get wet due to rains, they take these grains out into the sunlight to dry, and once these are dry, they take them back inside.

By more precise details for quantitative evaluation of number of processing elements (Ant agents) versus human brain neuronal cells. And by referring to the investigational analysis of observed cooperative behavioral social insects intelligent learning phenomenon among ants at ACS. The analogical comparative study of ants at ACS versus human brain's intelligence is performed by contributing group of neuronal cells inside the brain. This comparison characterized by its interdisciplinary, and challenging educational issue resulting in the very interesting quantitative findings briefly presented as follows. The experts estimate that an ant brain contains about 250,000 brain cells [7]. That number pales in comparison to the human brain, which is believed to contain over 86 billion neurons. However, for the ant, its brain is quite powerful. A human brain has about 10,000 million so a colony of 40,000 ants has collectively the same size brain as a human. However, in natural real world, some of ant colonies may be composed of millions of ant agents (members) [1]. Furthermore, by referring to [8], therein announced in words that : " of all the insects in the world, the ant can claim to have the largest brain". In fact, some scientists believe the ant has the largest brain in proportion to its size of any creature on Earth!". Accordingly, Ants are widely considered to be the smartest insect in the world [7]. Moreover, in the analogical context of bidirectional interactive human's educational process in classrooms versus ACS performance. Scientists have announced the finding that ants observed to be the first example of nonhuman animal that are capable of performing two-ways interactive instruction.[9].

This interdisciplinary research article motivated by Ant Colony's behavioral observations for two optimal selective cooperative decisions made by all of colony's ant-mates. Firstly, the decision related to

considering optimal selectivity problem between two food sources, one of them contains higher sugar amount than the other source. Both sources are simultaneously sited equidistant away from the original nest' colony site [10]. Herein, the analysis and evaluation of the two adopted modeling problems presented considered to be involved by self-organized mechanism. That's pointing out to perform optimal convergence of tandem running process directing towards the either better source with trailing marked by higher pheromone (for the first problem). Or directing the way to -even distant best nest site (for the second problem). In more details, the ants recruited as foragers visited the source containing higher amount of sugar which marked by the trailing pheromone resulting in directing towards the source with greater sugar amounts rather than visiting the source with lower in sugar [10]. Secondly, on the other hand, ACS performed optimal selectivity decision for finding (best) sheltered nest. That is by following stimulation of released a pheromone from her mandibular gland to avoid its vulnerability to presumable danger. Interestingly, it is noticed that selectivity case is greatly similar to optimal selection process of food source. This problem is following the tandem running regulation helping to carry over directions to new nests. That in accordance with what has been announced by O'Shea-Wheller [11]: "Colonies counteract the difficulty of finding a distant nest, simply by increasing the rate at which individuals give each other 'directions'". And thus, the amount of directional information that a colony gathers increases as a function of migration distance, sort of like a self-organizing route planner". Specifically, the two problematic issues achieve optimality of decisions by modulating the rate of 'tandem running', in which ants workers teach each other the route to either a better food source or a new nest site. In brief, both of the suggested problems are autonomously (Self-organizing) perform selective searching considering speed-accuracy trade off for optimum decision to reach either best source or nest site. Recently, three research papers have been published which concerned with comparative analogical study for both above problems originated from ants' swarm intelligence after application of Tandem Running Technique, which published at [12][13][14].

Herein, the intelligent behavioral performance of the second nonhuman animal (mouse), is demonstrated. Noticed the number of neuronal cells functioned as place field cells (at hippocampus mouse's brain area). That number is in well correspondence with the number of ant agents recruited for either ACS decision makers in the above [10][11]. More precisely, the contribution of neurons as place field cells at the mouse's hippocampus area resulted in minimum value of error by reaching (Cramer-Rao limit), while solving reconstruction problem [15][16][17][18]. Recent studies on the hippocampus have led to a new model of how the hippocampus forms and replays memories and so, the mouse had to run back and forth through the maze, from left to center to right, back to center, then left, and so on. The critical choice point was as the mouse left the center arm and had to remember whether it had most recently come from the left or right arm and choose the alternate direction [19][20][21]. Therefore, both introduced issues given at this paper's abstract, seem to be tightly related to the interactive learning responses of the two suggested types of nonhuman creatures (Ant & Mouse) with the natural environment they are living in [22].

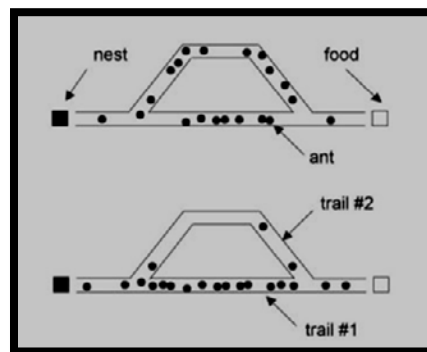
At [12], the paper dealt with the Comparison for the performance of some Behavioral Learning system models versus of an Ant Colony Optimization performance. The comparison between the above two problems originated from ants' swarm intelligence behavioral performance after application of Tandem Running Technique, is presented at [13]. The rest of this paper is organized in five sections, in addition to

the above previous first introductory section. These sections are briefly given as follows. The second next section presents revising of the selectivity decisional process of an optimal (best) food source. A review of ant colony migration strategies from the home nest to another non-vulnerable one to any potential danger is given at the fourth section. At the fourth section, analysis of reconstruction problem solution by a mouse inside Figure eight (8) maze is presented. The obtained simulation results are given at the fifth section. Finally, some interesting conclusive remarks are introduced at the last sixth section.

## 2 Ant's Selection of Best Food Source

### 4.1 Selection Between Two Pathways Using Pheromones' Marking

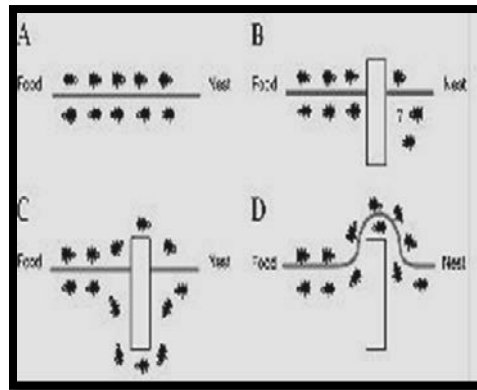
If an experimenter offers a colony of mass recruiting ants, one of the species using pheromones to food trails, two food sources simultaneously and at equal distances from the nest, but one is higher in sugar content than the other, most of the foragers will usually go to the source higher in sugar. Some of the foragers will feed from the source lower in sugar, but on the average, their numbers will be much lower than those going to the better source. This is of course a good decision for the survival and reproduction of the colony: the ants concentrate on the food source that provides the most calories with the least amount of effort. But how do they do this? How do they “decide” which source is better and how do they coordinate their efforts so as to exploit it preferentially?.[10]. Herein, a proposed explanation is currently introduced for how towards the better source containing the greatest amounts of pheromone and differentially choose these trails over those to lesser sources marked with smaller amounts of pheromone, illustration of how a group of ants able to optimally decide the selectivity of a shorter path to reach the food source via an asymmetrical simplified bifurcations network (referring to Fig.1.).



**Figure 1 Schematic illustration of the ant algorithm. (At the top). Selection of a shorter path between a nest and a food source by natural ants. The ants travel between the nest and food through trail #1 and trail #2. Initially, ants are distributed equally on both trails. (At the bottom) . Since trail #1 is shorter than trail #2, trail #1 becomes their favorite pathway with a higher pheromone concentration. (Adapted from[23])**

### 4.2 Selection of Minimum Pathway Between Source and Nest

Referring to Fig.2, in the case of bifurcation occurrence due to an existence of an obstacle at some point placed on the pathway between the nest site and that of the source, the transportation process of food (from food source) to food store (nest). is illustrated behavioral ants' responses shown at the simplified sketched figure considering the pheromone trail between nest and food source.



**Figure 2. Ant Behavior A. Ants in a pheromone trail between nest and food; B. An obstacle interrupts the trail; C. Ants find two paths to go around the obstacle; D. A new pheromone trail is formed along the shorter path.**

Accordingly, the persistent or recurring food sources may also be available, such as the aphids or scale insects 'farmed' by many ant species. The best strategy is often to remember rewarding foraging sites but also to be flexible enough to exploit newly discovered food and to select the better sources from those available. To this end, information directing nest mates to food also enables them to select the highest quality food find when multiple resources are available [10].

### **4.3 Selection of one Path Between Two Diamond's Branches**

Referring to [24], it is announced that, one of the most striking features of an ant colony's behavior is its capacity for the spatial organization of foraging activity. The use of trail pheromone to guide fellow workers in the nest to a large food source or rich foraging zone has been extensively studied [25] and obviously contributes to foraging efficiency. We have recently, however, been able to show that trail laying and trail following behavior are more than just a means of communicating a food source's location. When more than one trail is present at a time, the interactions between foragers and the trails can lead to the collective selection of the shortest path or the best food source, despite the fact that individual foragers have no means of making such choices. Referring to the published research work at [26], it reveals the importance of the shape of trail networks for foraging in ants and emphasizes the underestimated role of the geometrical properties of transportation networks in general.

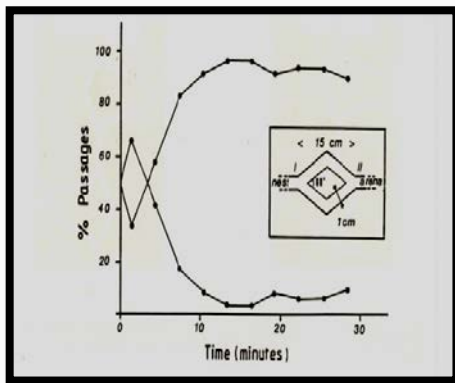


Figure 3. Percentage of the ant workers per three minutes period that are passing on the two branches of the diamond bridge (Adapted from [27])

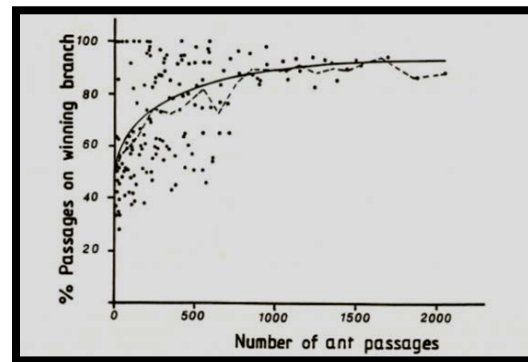


Figure 4. Percentage of the ant workers that are passing on the collectivity selected winning branch of the diamond bridge (shown at Fig.3) resulting after 20 experiments measured every three minutes(dots). Noting that horizontal axis represents the cumulative number of ant passages on both branches of the diamond bridge. The dashed curve represents the average values. The solid curve represents the average of 200 monte-Carlo Simulations.(Adapted from[24]).

## 5 Organization of Colony Migration

This section composed of three subsections which concerned with recruitment of signals used by successful foragers or nest site scouts, but another fundamental type of communication is alarm signaling. In social insects [28], defensive behavior is closely connected with alarm signals that either recruit nest mates to combat a potential-danger or warn them to stay away [29][30][31].

### 6.1 A. Application of Tandem Running Approach for performing ACS Emigration Function

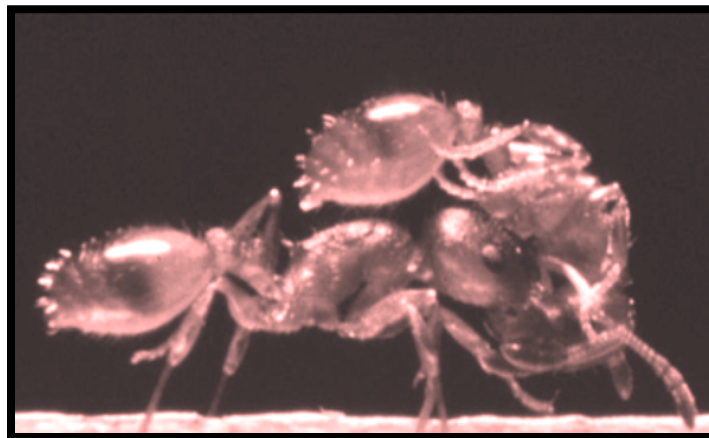
By referring to [32], detailed of the emigration functions is organized by a minority of active scouts, roughly *one-third of the colony's workers*. Each of these scouts sets out from the damaged nest to find a new home, thoroughly inspecting any candidate that she finds. If it passes muster, she returns to the old nest to reaching the site, but even these broken tandems recruit ants, because the orphaned follower enjoys a higher chance than a naive searcher of finding the target. Tandem followers make their own assessment of the site and may also begin to recruit. *The resulting positive feedback increases the site's population*, until it reaches a critical level and triggers a dramatic change in behavior She uses a behavior called tandem running, in which she attracts a single recruit to follow her toward the new site as shown at Figure 5. Their progress is slow and halting, as the leader must pause frequently to allow her follower to catch up. The pair often lose contact for good before reaching the site, but even these broken tandems recruit ants, because the orphaned follower enjoys a higher chance than a naive searcher of finding the target. Tandem followers make their own assessment of the site and may also begin to recruit. The resulting positive feedback increases the site's population until it reaches a critical level and triggers a dramatic change in behavior. Scouts cease tandem runs from the old nest, and instead begin to carry nest mates, one at a time, to the new site Figure 6. *Scouts cease tandem runs from the old nest, and instead begin to carry nest mates, one at a time, to the new site. These transports are roughly three times faster than*



*tandem runs, and population growth accelerates sharply*. Over the next few hours, the entire colony is brought to its new home. inform other scouts of its location.



**Figure 5.** Worker ants teach others the way to food with a poking and prodding technique called "tandem running." These ants have been daubed with paint for tracking purposes. Additionally, by adopting tandem running, a single ant is slowly led to a candidate site performing the migratory behavior forms an intrinsic part of the life histories the house hunting ant *Temnothorax albipennis*. colonies. [Adapted from [32] and image from Nigel Franks and Tom Richardson / University of Bristol]



**Figure 6.**Recruitment behavior used in emigration by *Temnothorax albipennis* colonies. Social transport, in which a single nest mate is rapidly carried to the new site. (Adapted from [32])

## 6.2 Selection Between Two Target Nests

Referring to [28], the migratory behavior forms an intrinsic part of the life histories of many organisms but is often a high-risk process. Consequently, varied strategies have evolved to negate such risks, but empirical data relating to their functioning are limited. In this study, we use the model system of the house hunting ant *Temnothorax albipennis* to demonstrate a key strategy that can shorten migration exposure times in a group of social insects. Colonies of these ants frequently migrate to new nest sites, and due to the nature of their habitat, the distances over which they do so are variable, leading to fluctuating potential costs dependent on migration parameters. Regarding to the closest resemblance of ants' lifestyle with respect to that of human beings. local advanced intercommunication observed via chatting

signaling among colonies' agents (ants). These signaling findings have been basically originated in accordance with the intercommunicative distributed collective cognition decisions among ants' colony members. Specifically, in the case of anticipated danger analysis, evaluation and deciphering of emerged alarm communication signals against predators resulting in variety of behavioral responses. In more details, considering the colony of *Temnothorax rugatulus*, it has been reported that alarm signaling pheromone while electing two different behaviors is dependable upon different context.

### 6.3 Binary Choice Between Two Nests [28]

Referring to Fig.5, When an ant was tethered inside an unfamiliar nest site and unable to move freely, she released a pheromone from her mandibular gland that signaled other ants to reject this nest as a potential new home, presumably to avoid potential danger. Accordingly, it is clearly possible that this pheromone's function can improve an emigrating behavioral response for colony's nest site selection performance. Colonies were given a binary choice between a nest with tethered ants and a nest that had five strings but no ants. By referring to Figure 7, concerned with test arena, were given a binary choice between a nest with tethered ants and a nest that had five strings but no ants. These two target nests were first placed adjacent to one another against one wall of the test arena. The home nest containing the colony from which the tethered ants were taken was then placed against the center of the wall opposite to the location of the target nests. Finally, the roof of the home nest was removed to induce migration.

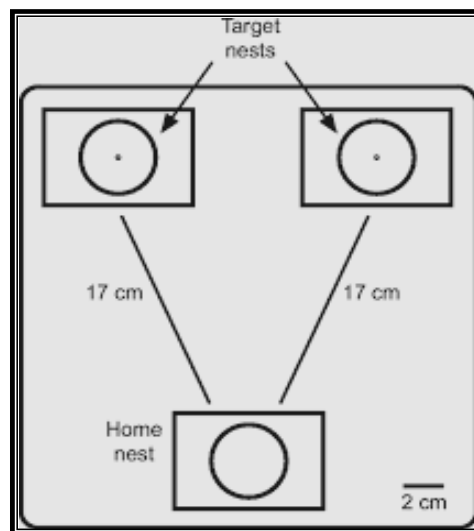


Figure 7. Experimental arena for nest choice tests. Colonies initially lived in the home nest, from which the roof was removed to induce migration. Colonies were allowed to choose between two target nests, which were identical in design but contained different materials (see Materials and methods for details). The arena size was 20×20 cm and 1 cm in height.(Adapted from [28]).

## 7 Reconstruction Problem Solving by a Mouse

### 8.1 Function of Brain's Hippocampus Area [16]

Hippocampus is said to be involved in “navigation” and “memory” as if these were distinct functions. In this issue of *Neuron* [17]. It provides evidence that the hippocampus retrieves spatial sequences in support of memory, strengthening a convergence between the two perspectives on hippocampal function. The following four interesting findings presenting basic characteristics aspects of the Brain's Hippocampus Area, that obtained after a set of distinct research works published recently to support

investigational research inside Fig.8 maze, aiming to shed light on the function of brain 's hippocampus Area:

a) Referring to [21], experimental testing performed for hippocampus brain area observed neural activity results in very interesting findings. Therein, ensemble recordings of 73 to 148 mouse's hippocampal neurons were used to predict accurately the animals' movement through their environment, which confirms that the hippocampus transmits an ensemble code for location. In a novel space, the ensemble code was initially less robust but improved rapidly with exploration. During this period, the activity of many inhibitory cells was suppressed, which suggests that new spatial information creates conditions in the hippocampal circuitry that are conducive to the synaptic modification presumed to be involved in learning. Development of a new population code for a novel environment did not substantially alter the code for a familiar one, which suggests that the interference between the two spatial representations was very small. The parallel recording methods outlined here make possible the study of the dynamics of neuronal interactions during unique behavioral events.

b) The hippocampus is said to be involved in "navigation" and "memory" as if these were distinct functions [17]. In this issue of *Neuron* this research paper evidence has been provided that the hippocampus retrieves spatial sequences in support of memory, strengthening a convergence between the two perspectives on hippocampal function.

c) Recent studies have reported the existence of hippocampal "time cells," neurons that fire at particular moments during periods when behavior and location are relatively constant as introduced at [18]. However, an alternative explanation of apparent time coding is that hippocampal neurons "path integrate" to encode the distance an animal has traveled. Here, we examined hippocampal neuronal firing patterns as rats ran in place on a treadmill, thus "clamping" behavior and location, while we varied the treadmill speed to distinguish time elapsed from distance traveled. Hippocampal neurons were strongly influenced by time and distance, and less so by minor variations in location. Furthermore, the activity of different neurons reflected integration over time and distance to varying extents, with most neurons strongly influenced by both factors and some significantly influenced by only time or distance. Thus, hippocampal neuronal networks captured both the organization of time and distance in a situation where these dimensions dominated an ongoing experience as illustrated at Figure 8. in below. [18].

d) Referring to Figure 9., to receive reward at the ends of maze arms, the rat had to run back and forth through the maze, from left to center to right, back to center, then left, and so on. The critical choice point was as the rat left the center arm and had to remember whether it had most recently come from the left or right arm and choose the alternate direction. The main finding in this study is that, as animals approached the critical choice point, the amount of hippocampal replay of both left and right routes predicted accurate choices. The optimal selected right choice is obtained in accordance with the stored experience at synaptic connectivity after its interaction with its environment [19][22]. Therefore, the observed optimal selective moving direction of the mouse is reached as a spontaneous response towards being inside Figure-of-eight-(8)maze.



Figure 8. Dissociation between elapsed time and path integration in the hippocampus during the delay period of a working memory task required the mouse to run on a treadmill for either a fixed amount, (Adapted from [18]).

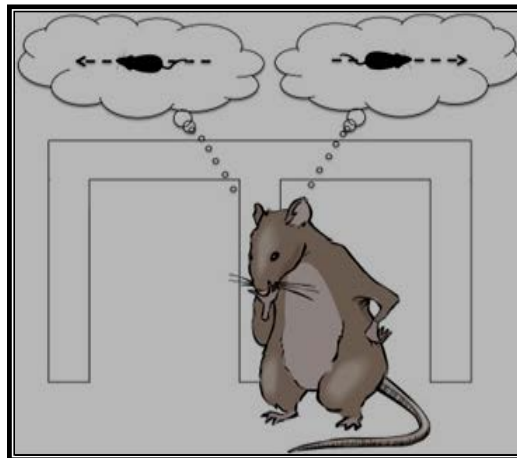


Figure 9. Retrieving memories of mouse's making decisions to select the correct direction in accordance with the stored experience at synaptic connectivity pattern of mouse's hippocampus area. (Adapted from [19]).

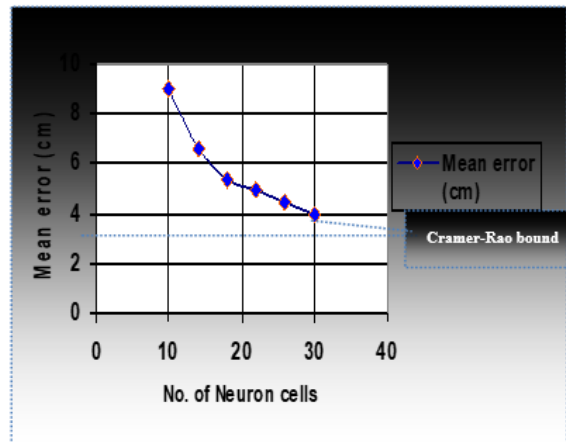
## 8.2 Maze Reconstruction Problem

Referring to measured mean error results shown at Table1, for solving reconstruction (pattern recognition) problem by a mouse inside a figure of eight (8) maze [15][16][20] That solution is carried out in accordance with the sliding of time window forward, the entire time course of  $x$  can be reconstructed from the time varying-activity of the neural population. Furthermore, that table is graphically represented as the algorithmic performance learning curve referred to Figure 11. It is noticed that decreasing of the mean error values is similar as an exponentially decayed function converged to some limit value versus (place field) cells. That converged limit (by the increased number of cells) reaches some value, named as Crammer-Rao bound. Originally that limiting bound is deduced from Fisher's information given in tabulated results (Table.1), obtained after an experiment for trials to solve reconstruction process from a pattern recognition problem [15]. Interestingly, the effect of increasing number of place field neurons in rat's brain hippocampus is analogous to the increased iterative training trials observed by the behavioral performance of some other nonhuman animals such as dogs and cats. These behaviors observed as results of performed Pavlov's and Thorndike's psycho-experimental works respectively.[33][34][35][36][37], in addition to the introduced comparative analogy of quantified

learning creativity in humans versus behavioral learning performance in animals: Cats, Dogs, Ants, and Mice [38].

**Table 1 Relation between Number of Cells and Mean Error in Solving Reconstruction Problem**

# neuron cells	10	14	18	22	26	30
Mean error (cm)	9	6.6	5.4	5	4.5	4



**Figure 10..The dashed line indicates the approach to Cramer-Rao bound based on Fisher information adapted from [15]**

## 9 Simulation Results

The presented simulation results have been in agreement with and supported by some obtained realistic computer modeling results, and announced findings published at a set of pieces of research manuscripts [39],[40],[41],[42],[43],[44], and [45]. Referring to Figure 4., the performance of an Ant Colony via the winning branch reaches some percentage value [%] after passing collectively a group of ant workers. This performance is dependable upon individual differences of ant colonies. This differences are simulated by Fig.9 while reaching the solution of Travelling Salesman Problem with different values of intercommunication parameters among ant colony agents. Referring to Fig.10, the different learning rate values  $\eta = 0.05, 0.1, \text{ and } 0.2$  are analogous to different intercommunication values among ant agents. That is by considering self-organized responses. resulted in convergence to different relative error values  $e(n)$  percentage values [%], after any fixed time period {number of trials (n)}. Moreover, the three curves at Fig.11, are analogous to the results shown at Fig.4. Referring to Fig.12, it simulates the performance of mouse's convergence to bounding limit (of minimum error value) fixed Cramer Rao bound (Limiting value) at Fig.8. Moreover, Fig.12, simulates the individual difference performance of animals as that illustrated recently at [16], where the neuronal population activity is evaluated during solving of reconstruction problem inside Fig.(8) maze via application of hippocampus place cells.



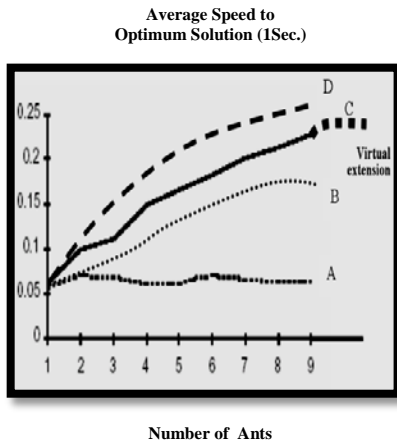


Figure 11.. Communication determines a synergistic effect with different inter communication levels among agents leads to different values of average speed

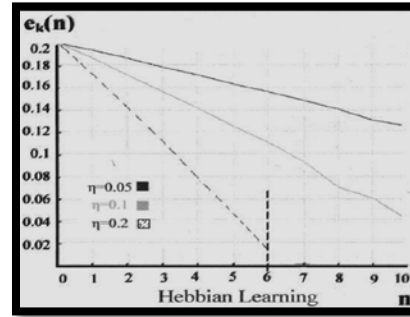


Figure 12. Adaptability performance concerned with self-organized Hebbian learning algorithm with learning rates (0.05,0.1,0.2). (Adapted from[46]).

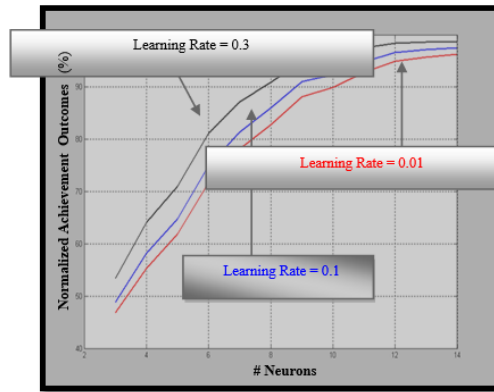


Figure 13. Illustrated the simulated output results presented as percentage degree[%]of normalized achievement outcomes versus # Neurons for different learning rate values  $\eta$  (0.01,0.1,and 0.3). and constant gain factor = 1

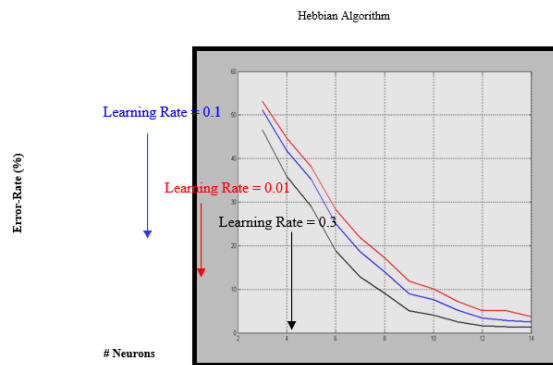


Figure 14. Illustrate learning performance error-rate with different learning rate values  $\eta$  (0.01,0.1,and 0.3). when #cycles = 300 and gain factor = 1.

## 10 Conclusions

This piece of research comes to four interesting conclusive remarks presented as follows:

- The existence of an obstacle at some point of ants' pathway (Figure 2), results in various trailing pheromone tracing. Due to asymmetry of obstacles' shape, the time needed to find the shorter pathway is directly proportional to the discovery of the minimum path.
- Ant colony optimization algorithms have been applied to many combinatorial optimization problems, ranging from quadratic assignment to fold protein or routing vehicles and a lot of derived methods have been adapted to dynamic problems in real variables, stochastic problems, multi-targets and parallel implementations. It has also been used to produce near-optimal solutions to the travelling salesman problem.
- The humble ant is not only capable of solving difficult mathematical problems, but is even able to do what few computer algorithms can - adapt the optimal solution to fit a changing problem, deepen our understanding of how even simple animals can overcome complex and dynamic problems in nature, and will help computer scientists develop even better software to solve logistical problems and maximize efficiency in many human industries.

## REFERENCE

- [1]. Ted R. Schultz " In search of ant ancestors Proc Natl Acad Sci U S A. 2000 Dec 19; 97(26): 14028–14029. Published online 2000 Dec 5 at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC34089/> doi: [10.1073/pnas.011513798](https://doi.org/10.1073/pnas.011513798)
- [2]. "Ants: Fun Facts about Ants & Ant Information for Kids - Pest World for Kids". Available online at: <https://pestworldforkids.org/pest-guide/ants/>
- [3]. E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press US, 1999.
- [4]. M. Dorigo and T. Stützle. *Ant Colony Optimization*. MIT Press, 2004.
- [5]. J. Kennedy, R. C. Eberhart, and Y. Shi. *Swarm Intelligence*. Morgan Kaufmann, 2001.
- [6]. How ants communicate? Available on line at <http://www.youtube.com/watch?v=gcHt5n3NGK0>  
**Uploaded on Jul 28, 2011.**
- [7]. *Seid, M. A.; Castillo, A.; Wcislo, W. T. (2011). "The Allometry of Brain Miniaturization in Ants". Brain, Behavior and Evolution. 77 (1): 5–13. Available online at: https://www.karger.com/Article/Abstract/322530.*
- [8]. John, and Sarah "Interesting Facts About Ants" Published at Free Materials (C) 1996. Available online at: <http://www.lingolex.com/jstefl.htm>

- [9]. Bjorn Carey "Ants help each other as teachers and pupils" Available Online by date Jan 11, 2006 at: [http://www.nbcnews.com/id/10806078/ns/technology\\_and\\_science-science/t/ants-help-each-other-teachers-pupils/#.Ws5hLS5ubIU](http://www.nbcnews.com/id/10806078/ns/technology_and_science-science/t/ants-help-each-other-teachers-pupils/#.Ws5hLS5ubIU) , © 2012 Live Science.com. All rights reserved.
- [10]. Ivan D. Chase, Abhijit V. Deshmukh & Naga Krothapalli:" How do Ants Decide Between Food Sources of Different Values? An evaluation of the Current Explanation and Associated Mathematical Models" Published at the PROCEEDINGS of the 2nd International Workshop on the Mathematics and Algorithms of Social Insects Georgia Institute of Technology, Atlanta, GA.30332, December 15–17, 2003, pp. 41-46.
- [11]. O'Shea-Wheller, T. A. *et al.* (2016). Migration control: a distance compensation strategy in ants, *The Science of Nature*, DOI 10.1007/s00114-016-1386.
- [12]. Hassan M. H. Mustafa, Fadhel Ben Tourkia, Ramadan Mohamed Ramadan "On Analysis and Evaluation of Comparative Performance for Selected Behavioral Neural Learning Models versus One Bio-Inspired Non-Neural Clever Model (Neural Networks Approach) Open Access Library Journal, Vol.3 No.10, October 31, 2016.
- [13]. Hassan M. H. Mustafa, and Fadhel Ben Tourkia "On Comparative Analysis and Evaluation Of Social Insect Colonies' Behavior During Exploring Food Sources and Their Migration to A New Nest Versus Two of Neural Networks' Learning Paradigms. (Tandem Running Approach)" Published Journal IJATTMAS volume III issue XI. Nov. 2017 Page 33-41.
- [14]. Hassan M. H. Mustafa, Fadhel Ben Tourkia. "On Application of Neural Networks' Modeling for Analytical Comparative Study between Two Optimally Selected Made Decisions by Ant Colony Systems". *American Journal of Educational Research*. 2018; 6(4):308-318. doi: 10.12691/education-6-4-3.
- [15]. Zhang , K. , Genzburg ,I. ,and Sejnowski,T.J. , 1998 "Interpreting neuronal population activity by reconstruction" *Journal of Neurophysiology*, 79:1017- 44,1998.
- [16]. Zhang, Iris Ginzburg, Bruce L. Mcnaughton, and Terrence J. Sejnowski "Interpreting Neuronal Population Activity by Reconstruction: Unified Framework With Application to Hippocampal Place Cells". Downloaded from <http://jn.physiology.org/> by 10.220.33.3 on October 29, 2016.
- [17]. Singer et al. (2013) "Hippocampus: Remembering the Choices" published at *Neuron*. Mar 20, 2013; 77(6): 999–1001. Available Online-at: [http://www.researchgate.net/publication/236073863\\_Hippocampus\\_remembering\\_the\\_choices](http://www.researchgate.net/publication/236073863_Hippocampus_remembering_the_choices)
- [18]. Kraus BJ<sup>1</sup>, Robinson RJ 2nd, White JA, Eichenbaum H, Hasselmo ME "Hippocampal "time cells": time versus path integration" *.Neuron.* 2013 Jun 19;78(6):1090-101. doi: 10.1016/j.neuron.2013.04.015. E pub 2013 May 23. Available online at <http://www.ncbi.nlm.nih.gov/pubmed/23707613>

- [19]. Howard Eichenbaum "Hippocampus: Remembering the Choices" *Neuron*, Volume 77, Issue 6, p999–1001, 20 March 2013.
- [20]. Sejnowski, T.J., 1999: "Neural pulse coding" foreword article for (Pulsed neural networks), MIT press, 1999, pp 13-23.
- [21]. Wilson, M. A. and McNaughton, B. L., Dynamics of the hippocampal ensemble code for space. *Science*. 1993 Aug 20; 261(5124):1055-8. Available online at <http://www.ncbi.nlm.nih.gov/pubmed/8351520>
- [22]. Fukaya, M., et al. Two level Neural Networks: Learning by Interaction with Environment, 1st ICNN, San Diego, 1988.
- [23]. Yunlong Liu and Hiroki Yokota "Artificial ants deposit pheromone to search for regulatory DNA elements". Available online at: <https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-7-221>. **Published:** 30 August 2006. The image available online at: [http://media.springernature.com/full/springer-static/image/art:10.1186/1471-2164-7-221/MediaObjects/12864\\_2006\\_Article\\_604\\_Fig1\\_HTML.jpg](http://media.springernature.com/full/springer-static/image/art:10.1186/1471-2164-7-221/MediaObjects/12864_2006_Article_604_Fig1_HTML.jpg)
- [24]. S. Goss, R. Beckers, J. L. Deneubourg, S. Aron, J. M. Pasteels "How Trail Laying and Trail Following can Solve Foraging Problems For Ant Colonies" *Behavioral Mechanisms of Food Selection* pp 661-678. NATO ASI Series, Vol. G 20 Behavioural Mechanisms of Food Selection Edited by R. N. Hughes © Springer-Verlag Berlin Heidelberg 1990. Available online at: [https://link.springer.com/content/pdf/10.1007/978-3-642-75118-9\\_32.pdf](https://link.springer.com/content/pdf/10.1007/978-3-642-75118-9_32.pdf)
- [25]. Wilson EO (1971). *The insect societies*. Harvard University Press, Cambridge Massachusetts
- [26]. Simon Garnier, Maud Combe, Christian Jost, and Guy Theraulaz "Do Ants Need to Estimate the Geometrical Properties of Trail Bifurcations to Find an Efficient Route? A Swarm Robotics Test Bed". Published: on March 28, 2013.
- [27]. Deneubourg JL, Aron S, Goss S, Pasteels JM. (1989a) The self-organizing exploratory pattern of the Argentine ant. *J Ins Behav* in press.
- [28]. Sasaki, Bert Hölldobler, Jocelyn G. Millar, Stephen C. Pratt "A context-dependent alarm signal in the ant *Temnothorax rugatulus*". Published at the *Journal of Experimental Biology* 2014 217: 3229-3236; doi: 10.1242/jeb.106849. Available online at: <http://jeb.biologists.org/content/217/18/3229>.
- [29]. Blum, M. S. (1969). Alarm pheromones. *Annu. Rev. Entomol.* 14, 57-80.

- [30]. Blum, M. S. (1985). Alarm pheromones. In *Comprehensive Insect Physiology, Biochemistry and Pharmacology: Behaviour*, Vol. 9 (ed. G. A. Kerkut and L. I. Gilbert), pp. 193-224. New York, NY: Pergamon Press.
- [31]. Crewe, R. M. and Fletcher, D. (1974). Ponerine ant secretions: the mandibular gland secretion of *Paltothyreus tarsatus* Fabr. *J. Entomol. Soc. South Africa* **37**, 291-298.
- [32]. Pratt S.C. (2010) Nest Site Choice in Social Insects. In: Breed M.D. and Moore J., (eds.) *Encyclopedia of Animal Behavior*, volume 2, pp. 534-540 Oxford: Academic Press. Available online at: <http://www.elsevier.com/locate/permissionusematerial>
- [33]. Pavlov, I.P. *Conditional Reflex, An Investigation of The Psychological Activity of the Cerebral Cortex*, New York, Oxford University press, 1927.
- [34]. Hampson, S.E. *Connectionistic Problem Solving, Computational Aspects of Biological Learning*, Berlin, Birkhouser, 1990.
- [35]. Thorndike E.L. *Animal Intelligence*, Darien, Ct. Hafner, 1911.
- [36]. Hassan H. and Watany M. On Mathematical Analysis of Pavlovian Conditioning Learning Process using Artificial Neural Network Model, 10th Mediterranean Electro technical Conf., May 29-31, 2000, Cyprus.
- [37]. H. M. Hassan , and M. Watany. "On Comparative Evaluation And Analogy for Pavlovian and Throndikian Psycho-Learning Experimental Processes Using Bioinformatics Modeling", published at AUEJ, 6,3, 424-432, July. 2003.
- [38]. Hassan, M.H., 2008 " A Comparative Analogy of Quantified Learning Creativity in Humans Versus Behavioral Learning Performance in Animals: Cats, Dogs, Ants, and Rats.(A Conceptual Overview), published at WSSEC08 conference held on 18-22 August 2008, Derry, Northern Ireland.
- [39]. H.M. Hassan, "On Mathematical Modeling of Cooperative E-Learning Performance During Face to Face Tutoring Sessions (Ant Colony System Approach)"published at IEEE EDUCON 2011,on Education Engineering–Learning Environments and Ecosystems in Engineering Education, held on April 4-6, 2011,Amman, Jordan. Available on line at: [http://www.google.com.sa/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=9&ved=0CHQQFjAI&url=http%3A%2F%2Feditlib.org%2Fd%2F45687&ei=rsUFU66SAc-c0wWakIHAAQ&usq=AFQjCNFXdog2WcQE\\_3DE5-8sVp7aaVH4Lw](http://www.google.com.sa/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=9&ved=0CHQQFjAI&url=http%3A%2F%2Feditlib.org%2Fd%2F45687&ei=rsUFU66SAc-c0wWakIHAAQ&usq=AFQjCNFXdog2WcQE_3DE5-8sVp7aaVH4Lw)
- [40]. H. M. Hassan. "On Learning Performance Evaluation for Some Psycho-Learning Experimental Work versus an Optimal Swarm Intelligent System.", Published at ISSPIT 2005 (18-20 Dec.2005). [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1577175&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D1577175](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1577175&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D1577175)



- [41]. H.M. Mustafa "A tutorial titled: Building up bridges for natural inspired computational models across behavioral brain functional phenomena; and open learning systems", that has been presented at the International Conference on Digital Information and Communication Technology and its Applications (DICTAP 2011) held at Universite de Bourgogne, Dijon, France. (June 21-23, 2011). Available online at: <http://dictap2011.sdiwc.us/tutorials.php>
  
- [42]. Hassan M. H. Mustafa, and Fadhel Ben Tourkia "On Analysis and Evaluation of Learning Creativity Quantification via Naturally Neural Networks' Simulation and Realistic Modeling of Swarm Intelligence" published at the proceeding of the conference Eminent Association of Researchers in Engineering & Technology(EARET).To be held on 8-9 January 2018.
  
- [43]. Hassan M. H., et.al"On Comparative Analogy between Ant Colony Systems and Neural Networks Considering Behavioural Learning Performance" Journal of Computer Sciences and Applications, 2015, Vol. 3, No. 3, 79-89 Available online at <http://pubs.sciepub.com/jcsa/3/3/4> © Science and Education Publishing DOI:10.12691/jcsa-3-3-4.
  
- [44]. Hassan M. H. "Analytical Comparison of Swarm Intelligence Optimization versus Behavioral Learning Concepts Adopted by Neural Networks (An Overview) American Journal of Educational Research<http://pubs.sciepub.com/education/3/7/2/index.html> Vol. 3, No. 7, 2015, pp 800-806. doi: 0.12691/education-3-7-2
  
- [45]. Hassan M. H., et.al "Comparative Performance Analysis and Evaluation for One Selected Behavioral Learning System versus an Ant Colony Optimization System" Published at the Proceedings of the Second International Conference on Electrical, Electronics, Computer Engineering and their Applications (EECEA2015), Manila, Philippines, on Feb. 12-14, 2015.
  
- [46]. Hassan M. H., et.al. "On Assessment of Brain Function Adaptability in Open Learning Systems Using Neural Networks Modeling (Cognitive Styles Approach). Journal of American Science 2011; 7(9): 238-247]. (ISSN: 1545-1003). <http://www.americanscience.org>

# Application of an Artificial Neural Network for Early and Accurate Diagnosis of Parkinson's disease

<sup>1</sup>Akshar Agarwal

<sup>1</sup>Hamden High School, Connecticut, United States;  
[aakshar2000@gmail.com](mailto:aakshar2000@gmail.com)

## ABSTRACT

Parkinson's Disease (PD) is one of the most common neurological disorders, affecting more than ten million people globally. The hallmark symptoms of PD are tremors, limb rigidity, and imbalance. PD shares many of these symptoms with other disorders, making it difficult to diagnose. Furthermore, due to the lack of definitive laboratory tests, PD is poorly diagnosed with subjective examinations such as family history evaluations, resulting in high misdiagnosis rates. Recent research shows that an additional symptom, dysphonia, is uniquely present in over 80% of PD patients. Dysphonia is a speaking disorder caused by involuntary muscle movement and other neurological factors in PD. In this project, that unique symptom. A cross-validated neural network was programmed to deliver rapid and accurate diagnoses using biomedical voice data from 195 patients of varying statuses. This automated, machine-learning based PD diagnostic tool was successfully created and functions with over 95% accuracy. This rate includes nearly zero false negatives and few false positives, showing significant improvement over previous attempts which had misdiagnosis rates of nearly 20%. A low probability of false negatives is favorable. The neural network was designed such that overfitting is avoided, and more features/data would further improve the algorithm's accuracy. An early and accurate diagnosis is critical for treating PD patients, and this project proposes a way to achieve that.

**Keywords:** Machine learning, Artificial intelligence, Parkinson's Disease, neural network, dysphonia

## 1 Introduction

Parkinson's Disease (PD) is one of the most common neurodegenerative disorders, affecting nearly 3% of adults aged 65 and older [2]. It is a common, yet highly complex disease which results from both genetic and environmental factors. PD has severe symptoms, most notably tremors, limb rigidity, and gait issues [2,3]. The disease is primarily characterized by the loss of dopaminergic neurons the midbrain [2]. Despite its prevalence, there currently exists no objective method of determining whether a patient is affected by PD. The diagnosis rests upon subjective evaluations such as family history evaluations, handwriting examinations, and response to medication. Although new imaging techniques have been developed to aid in the diagnosis of PD, there are still no proper diagnostic tools. Furthermore, many other diseases present with similar symptoms (MS, Fibromyalgia, Lewy-Body dementia), causing misdiagnosis rates to remain significantly high at approximately 24% [3]. Recent studies, including one performed by Kish et al, have found that drastic neuronal depletion has been observed by the time that motor symptoms emerge. Because dopaminergic loss occurs rapidly when left untreated, an early and accurate diagnosis of PD is essential [4]. An early and accurate diagnosis is essential to patient outlook, in terms of both health and

finances. Medication can aid in slowing neurological damage or controlling movement, and economically speaking, treatment costs grow exponentially with the severity of the disease. Recent research suggests that 80-95% of PD patients present with dysphonia, a speaking disorder, in the earliest stages [1]. Although the exact mechanism remains unknown, dysphonia is the result of various neurological and physical influences which occur with Parkinson's Disease. Because vocal impairment is one of the earliest indicators of PD, and voice samples can be collected non-invasively, utilizing voice-based diagnostic tools would provide a significant improvement over current diagnostic techniques. In this project, an artificial neural network was created to diagnose PD by utilizing voice samples of healthy and diseased patients.

## 2 Algorithm Development

The Statistics and Machine Learning Toolbox was used within MATLAB to develop an Artificial Neural Network which can classify patient condition given voice data.

### 2.1 Data

The dataset used here consists of 195 sustained vowel phonations from 31 patients, 23 of whom have been diagnosed (in various stages) with Parkinson's Disease. The patients' ages ranged from 46 to 85 years, as shown in Figure 2.1, and the time since diagnosis ranged from 0 to 28 years, as shown in Figure 1. The data was collected and provided by request from Dr. Max Little, in collaboration with the National Center for Voice and Speech (Denver, CO). The speech signals were recorded primarily using the Kay Pentax Multi-Dimensional Voice Program [1].

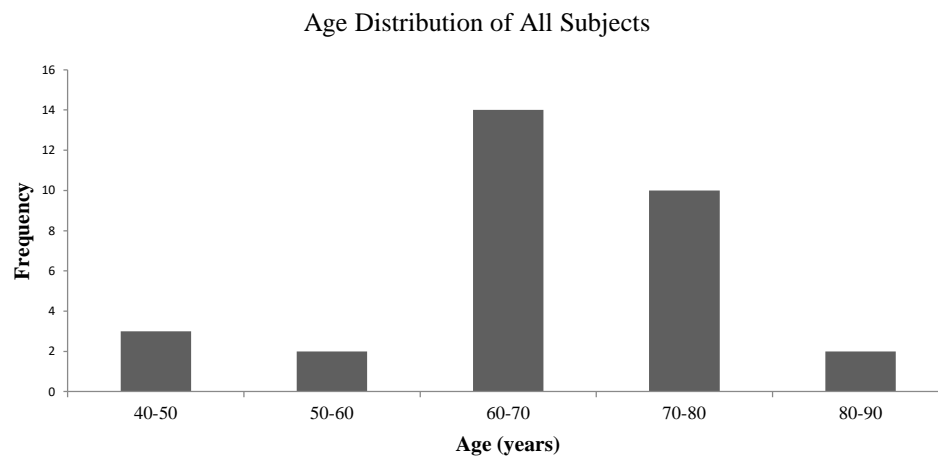


Figure 1: A histogram showing the distribution of subjects' ages.

#### 2.1.1 Features

The data consists of 22 features, including various measures of fundamental frequency, amplitude, et cetera. Other important features included jitter (relative changes in fundamental frequency), shimmer (relative changes in amplitude across vocal cycles), and several values which quantify the noise-to-harmonics ratios.

### 2.2 Neural Network

An artificial neural network (ANN) was designed here to classify subject condition as 1 (diseased) or 0 (healthy). A multilayer perceptron was created and validated with leave-one-out (LOO) cross-validation

and dropout layers. The LOO cross-validation entails training on n-1 trials and then testing on the remaining 1 trial for n iterations. Dropout entails selectively dropping or ignoring random neurons and their inputs. This is done to reduce both variance and bias in efforts of creating an optimal classifier which neither overfits nor underfits [6, 7]. Other basic architectural properties include a logistic sigmoidal transfer function between the hidden layer and output layer to force a binary output.

### 2.3 Bayesian Optimization

ANNs are commonly referred to as black box functions. There are many parameters whose influence on overall error is largely unpredictable. While it is not necessarily true that all ANNs operate as such, the model created here is certainly not convex. Its global minimum is difficult to locate, and thus Bayesian optimization, as described by Snoek et al. [5], was utilized for hyperparameter selection and model evaluation. Grid searches are computationally intensive and grow exponentially with more parameters. On the other hand, this method provides a more logical search and can be run with less resources. Bayesian optimization involves fitting a Gaussian process to the model to predict which parameters will return more optimal accuracies/errors. Figure 2 shows the resulting model of the objective function which was obtained by tracing outputs with two variable hyperparameters. This process calculates the expected improvement across parameters and seeks to minimize the value of the objective function (error).

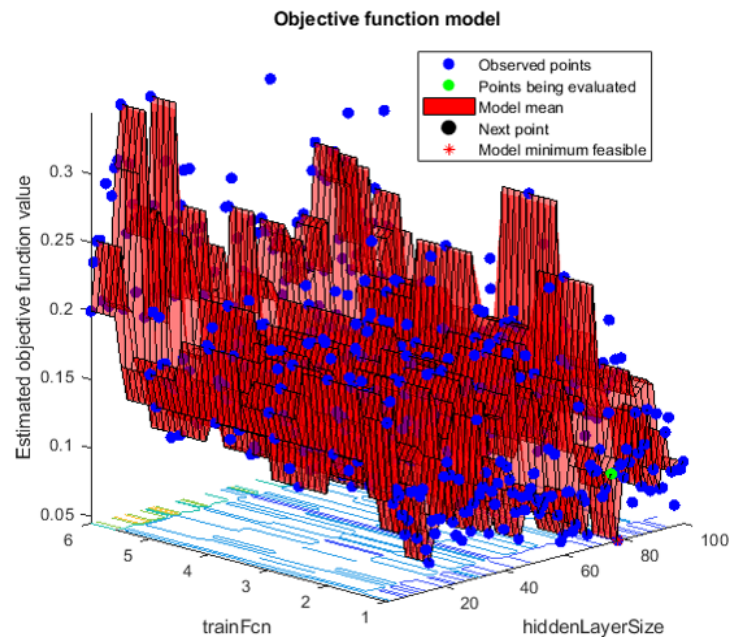


Figure 2: A model of the Artificial Neural Network obtained through the completion of one thousand iterations of Bayesian Optimization.

## 3 ANN Performance and Results

The optimal ANN architecture was determined to consist of a single hidden layer with 78 neurons. The transfer function between the input and hidden layers was the Elliot symmetric sigmoidal function, and the Levenberg-Marquardt algorithm was utilized for backpropagation.

$$x_{k+1} = x_k - [J^t J + \mu I]^{-1} J^t e \quad (1)$$

The Levenberg-Marquardt algorithm is one of the fastest and most accurate methods of backpropagation [5]. As seen in the equation above,  $J^t J$  represents the Hessian matrix, and  $J^t e$  is the gradient where  $J$  is the Jacobian matrix of first derivatives. The Hessian and Jacobian matrices describe the local curvature of a multivariate function.  $\mu$  is a scalar value which helps model the algorithm after Newton's method, which is particularly fast while approaching a global error minimum. This algorithm transforms from gradient descent to a quasi-Newton BFGS algorithm as  $\mu$  decreases.

With that architecture, the cross-validated artificial neural network yielded an overall accuracy of 95.9%, as seen in the confusion matrix in Figure 3.

**Parkinson's NN: Confusion Matrix**

NN Prediction	0 (Healthy)	42 21.5%	2 1.0%	95.5% 4.5%
	1 (PD)	6 3.1%	145 74.4%	96.0% 4.0%
		87.5% 12.5%	98.6% 1.4%	95.9% 4.1%
		0 (Healthy)	1 (PD)	
		Actual Condition		

**Figure 3: A confusion matrix displaying the accuracy of the ANN by class.**

A high probability of detection is the obvious objective, but a low false-negative rate is especially favorable in medical settings. The model created here functions with a false-negative rate of just 1.4%, indicating that patients with Parkinson's disease are not given an inaccurate diagnosis (healthy) at a statistically significant rate.

## 4 Conclusion

Parkinson's Disease is one of the most prevalent neurodegenerative disorders, and yet its complexity often leads to misdiagnosis and defers treatment for years. Detecting Parkinson's Disease in its early stages is crucial, and the nature of the disease does not allow for physicians to detect it efficiently. Misdiagnosis rates are significantly high, and this creates issues for patients both economically and

## REFERENCES

- [1] Little, M. A., McSharry, P. E., et al. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions on Bio-Medical Engineering*, 56(4), 1015.
- [2] Sewall G. K., Jiang J., Ford C. N. Clinical evaluation of Parkinson's-related dysphonia. *Laryngoscope*. 2006;116(10):1740–1744.



- [3] Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*. 1992;55(3):181-184.
- [4] Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease: pathophysiologic and clinical implications. *N Engl J Med*. 1988;318(14):876-880.
- [5] MathWorks, (2017). *Statistics and Machine Learning Toolbox: User's Guide (R2017b)*.
- [6] Srivastava, Nitish, et al. "Dropout: A simple way to prevent neural networks from overfitting." *The Journal of Machine Learning Research* 15.1 (2014): 1929-1958.
- [7] Kohavi, Ron. "A study of cross-validation and bootstrap for accuracy estimation and model selection." *Ijcai*. Vol. 14. No. 2. 1995.

# An Ocean of Opportunities in Artificial Human Optimization Field

**Satish Gajawada**

*Alumnus, Indian Institute of Technology Roorkee, INDIA*

*Founder, Artificial Human Optimization – A New Field*

[gajawadasatish@gmail.com](mailto:gajawadasatish@gmail.com)

“Great LEADERS don’t tell you what to do. They show you how it’s done” – From Internet

## ABSTRACT

Global Optimization Techniques like Genetic Algorithms, Particle Swarm Optimization, Ant Colony Optimization and other optimization techniques were used in literature to solve complex optimization problems. Many optimization algorithms were proposed in literature by taking the behavior of Birds, Ants, Fishes, Chromosomes etc. as inspiration. Recently, a new trend has begun in Evolutionary Computing Domain where optimization algorithms have been created by taking Human Behavior as inspiration. The focus of this paper is on optimization algorithms that were and are being created based on the behavior of Artificial Humans. In December 2016, a new field titled “Artificial Human Optimization” was proposed in literature. This paper is strongly meant to popularize “Artificial Human Optimization” field like never before by showing an Ocean of Opportunities that exists in this new and interesting area of research. A new field titled “Artificial Economics Optimization” is proposed at the end of paper.

**Keywords:** Artificial Intelligence, Machine Learning, Evolutionary Computing, Bio-Inspired Computing, Nature Inspired Computing, Genetic Algorithms, Particle Swarm Optimization, Ant Colony Optimization, Artificial Economics Optimization, Artificial Human Optimization

## 1 Introduction

Article [1] proposed a new field titled “Artificial Human Optimization (AHO)” by showing 13 abstracts of papers in this new and interesting area of research called as AHO field. Article [2] showed different reviews of experts for a particular work in AHO field. Article [3] proposed new optimization method titled “POSTDOC : The Human Optimization” which comes under AHO field. Article [4] popularized AHO field by “showing few opportunities in AHO field”, “proposing Multiple Strategy Human Optimization method”, “showing reviews for paper – POSTDOC: The Human Optimization”, “making corrections to earlier work”, “encouraging researchers to work in this new area” and “giving Artificial Human Optimization Award to researchers who worked in this new area”.

In this article, Section 2 shows reviews of papers in Artificial Human Optimization field. The reviews shown in article [4] are the reviews of a particular paper titled “POSTDOC: The Human Optimization”. In this article, reviews of other papers proposed in AHO field are shown.

Section 3 in this article shows many opportunities that exist in this new area of research. Section 4 shows “Doctor of Philosophy (PhD)” proposal in Artificial Human Optimization area.

In this article, Section 5 shows future work where a new field titled “Artificial Economics Optimization (AEO)” is defined. This section also shows some of the opportunities that exist in AEO area. Acknowledgements are shown in Section 6 of this Article.

## 2 Reviews

### 2.1 Review 1

We had a glance at your published article “POSTDOC : THE HUMAN OPTIMIZATION”. We found your article very **innovative, insightful and interesting**. We really value your **outstanding contribution** towards Scientific Community.

### 2.2 Review 2

The author have come up with a **new approach** of academia; Philosophy Of a Doctor (PoD) – A New Degree. I am not expert in this area and would suggest to send the work to some academic. I think this is not a suitable workshop for such work. It should be send to some academic workshop/conference/journal.

### 2.3 Review 3

Literature review: **very good**.

Experiments: None.

Overall a **good writing** but check if it is in the scope of the conference of PAKDD!

### 2.4 Review 4

The author, Satish Gajawada has proposed a new area Artificial Human Optimisation – An Introduction. The author has proposed an optimisation algorithm but there has been no application used to test the algorithm or the optimisation method on. More experiments are required. Also its not related to data mining and not in the scope of PAKDD or BDM.

### 2.5 Review 5

The author note" This article got recently published in “Transactions on Machine Learning and Artificial Intelligence”. I have copyrights with me. I am submitting this paper for re-publishing in your workshop in an attempt to popularize “Artificial Human Optimization” like never before." suggests that the work is already published so there is not point accept him for re-publication since the proceedings of this workshop will be published. **New and interesting area though**.

### 2.6 Review 6

We are very happy to inform you that your paper has been accepted (conditionally) for publication in journal. Our journal is a prestigious journal and it will be indexed in major indices ASAP. Please modify your paper based on reviewers comment, otherwise it will not be accepted for publication.

Type of paper : Research

Significance of the main idea(s) : Average

Originality : Average

Technical quality of the paper : Average

Awareness of related work : Average

Clarity of presentation : Average

Organization of the manuscript : Very Poor

References : Average

Paper Length : Average

How comfortable are you in reviewing this paper? : Confident

Overall comments and changes that MUST be made before Publication: The manuscript is not organized. The presentation is not clear.

Overall Recommendation: Marginally Accept.

## 2.7 Review 7

The Paper captioned “Ph.D: The Human Optimization” presented by the author has been reviewed in detail by the reviewers and found it a distinctive effort. Reviewers offer following remarks:

The paper is **strongly prepared** to provide technical grounds of the subject. The Introduction elucidates the essence of the proposed research. Besides, Literature Review carefully outlined with details focused on the current applicable methods.

The author intelligently developed the remaining part of the manuscript communicating the commitment and the pragmatic knowledge of the writer. The mathematical expressions were also strongly used to defend the current work. The methodology adopted showed proper evaluation and documentation. In this manner, it is acknowledged for publication as it is.

## 2.8 Review 8

The paper “Scientist: International Association of Artificial Human Optimization” developed generously that shows the technical grounds of the subject. The technical expertise has been greatly utilized in developing the paper under review. The Introduction endorsed systematically the extensive areas of the new study approach. Besides, Literature Review well explains and advocates the probe and gaps in this area of research. The author skillfully produced the rest of the manuscript exploring in depth this newly proposed area. Moreover, the optimization method is well devised and the reviewer’s comments on previous work are also commendable. In this manner, it is acknowledged for publication as it is.

## 2.9 Review 9

The Paper captioned “Hero: Transactions on Artificial Human Optimization” has been carefully reviewed by the authorities and declared it a well-composed paper. Reviewers offer following remarks:

The abstract focused on the rationale of the research in a logical way that seems to be a unique angles of the study. The Introduction of the paper well argued the true methodology of the research. Moreover, Literature review is carefully organized, covers the available methods with suitable details. Rest of the manuscript gradually covers the focused point of view that shows the applied knowledge of the authors. The paper also presents the relevant mathematical details with sufficient reference to the existing work.

The methodology adopted showed proper evaluation and documentation. Therefore, it is accepted for publication as it is.

### **2.10 Review 10**

The article contributes greatly on the areas of Artificial Intelligence. Researcher excellently worked on the area of study. Collectively, it is a great effort and the reviewers provided following comments about the manuscript. Abstract covers all the important aspects of the proposed methodology and well written in general. The essence of the presented approach is elaborated nicely in the Introduction section. Similarly, Literature review is organized well. Rest of the manuscript is also very well structured representing the dedication and knowledge of the researcher about the topic and skill on research. The manuscript shall be rated high on its technical quality. Therefore, manuscript shall be accepted for publication as it is.

### **2.11 Review 11**

The Paper captioned “Artificial Human Optimization –An Introduction” has been wisely appraised by the authorities and declared it a well-conceived paper. Reviewers offer following remarks:

The abstract previews the author’s approach and improved theories of the study that definitely yet to be explored and never produced before. The research procedure as described in the Introduction is exemplary. Moreover, Literature review is best framed, focus the information on the potential approaches with applicable facts. Rest of the manuscript gradually covers the ideal point of view that shows the applied knowledge of the authors. The paper also presents the relevant mathematical details with sufficient reference to the existing work. The methodology adopted showed proper evaluation and documentation. Therefore, it is accepted for publication as it is.

### **2.12 Review 12**

Information for the Contribution

1. Writing Skill and Quality (0-10): 8
2. Quality of content (0-10): 8
3. Fitness of title (0-10): 9
4. Significance for theory or practice (0-10): 9
5. Contribution and Originality (0-10): 9
6. Level of Innovation (0-10): 8
7. Quality of presentation (0-10): 8
8. Ripple effect to other authors (0-10): 10
9. Decisive overall recommendation (0-10): 9

In this paper the author clearly explains the research with effective method and good description. The topic is meaningful, and the research results are interesting to many specialized readers. So it's good work and acceptable.

Some more checkpoints for improving the quality of the final version (if need, author can pay attention to below checkpoints. It's not requirements but just comments for improving the paper)

- In abstract and Introduction, if need, the main goal can be say more clearly.
- The abstract can shows some information to understand the goal of this paper clearly, if need.
- The introduction can be emphasized the background and motivation more clear way.
- the results need reflect more clear explanation.
- Are there any not good phrases?, if so the authors need to pay attention to them.
- Are there any typos? If so, it need to be revised.
- Although paper is good from language point of view, but a little bit review towards sentences and/or grammars can turn it into beautiful paper.
- In conclusion, should elaborate why the study is important and the significance of the study more clearly.
- Overall, it's **very interesting** and the work proposed is useful, the paper is organized well and the presentation is clear.

### 3 Ocean of Opportunities

From previous section, it is clear that sme world-class experts have accepted the fact that Artificial Human Optimization is very interesting area of research. As shown in Article [4], there are also set of world-class experts who are against to the idea of Artificial Human Optimization.

Artificial Human Optimization field is in its early stage. The field got created in December 2016. There are literally so many opportunities in AHO field. New research papers can be published by replacing the optimization algorithm (like Genetic Algorithms, Particle Swarm Optimization, Ant Colony Optimization etc.) used in the published paper with algorithms in AHO field. Satish Gajawada et al. published nearly 10 papers using algorithms based on optimization algorithms like GA, PSO and DE. In this section titles of papers based on algorithms in AHO field are shown. The below given research paper titles are not published yet. The below titles of papers are future opportunities for researchers interested to work in AHO field:

- 1) Optimal Clustering Method Based on algorithms under AHO field [6].
- 2) Projected Clustering Using HPSO algorithm [7].
- 3) HBBO Based Projected Clustering Method [8].
- 4) Projected Clustering HPSO and Classification [9].
- 5) A Method of Initialization for HPSO Based Clustering Technique [10].
- 6) Supervised Projected Clustering Method Based on HBA algorithms [11].
- 7) Design optimization of non linear tapers for high power gyrotrons using hybrid Human Behavior Algorithms and space mapping methods [12].
- 8) A Semi-Supervised Projected Clustering Method Using HBBO [13].
- 9) A framework for classification using Human Behavior Algorithms based clustering [14].

HBBO, HBA, HPSO in above paper titles are algorithms under Artificial Human Optimization field. Details of these algorithms are given in Article [4]. A total of 15 papers under AHO field are shown in [4].

The above new project titles / research papers to be published are just research opportunities based on Satish Gajawada et al. authors. Imagine the number of new projects that are possible based on previous papers of crores of researchers across the globe. There are millions of papers possible because there are



so many papers published based on optimization algorithms like GA, PSO, DE and other optimization algorithms.

Besides new projects/papers there is possibility for so many associations, organizations and research labs across the globe. Few examples are shown below:

- 1) IIT Roorkee Artificial Human Optimization Lab
- 2) IEEE Artificial Human Optimization Society.
- 3) Indian Journal of Artificial Human Optimization.

There are so many possibilities to create new AHO associations, AHO journals, AHO conferences, AHO research labs and AHO societies etc.

Hence there is an Ocean of Opportunities in Artificial Human Optimization field.

#### **4 Doctor of Philosophy Proposal**

This proposal is meant for students who are interested to do “Doctor of Philosophy (PhD)” in Artificial Human Optimization.

“Optimization Algorithms based on Human behavior” is the title of the work.

Article [5] proposed HPSO algorithm which is the inspiration of this PhD proposal.

The description of HPSO algorithm is taken from Article [5] as it is and shown below in double quotes:

“HPSO is modified PSO, based on human behavior, which is proposed to improve the performance. In PSO, all particles only learn from the best particles Pbest and Gbest. Obviously, it is an ideal social condition. However, considering the human behavior, there exist some people who have bad habits or behaviors around us, at the same time, as we all known that these bad habits or behaviors will bring some effects on people around them. If we take warning from these bad habits or behaviors, it is beneficial to us. Conversely, if we learn from these bad habits or behaviors, it is harmful to us. Therefore, we must give an objective and rational view on these bad habits or behavior. In HPSO, we introduce the global worst particle, who is of the worst fitness in the entire population at each iteration. It is denoted as Gworst.”

We can add human behavior to existing algorithms like Genetic Algorithms, Ant Colony Optimization etc. and create algorithms like “Human Behavior Inspired Genetic Algorithms”.

In HPSO, Worst particle is introduced. Similarly we can add human behavior to existing algorithms. We can then see results without adding human behavior and results after adding Human behavior.

#### **5 Future Work**

“Artificial Economics Optimization” is a new field where algorithms are created by taking “Economics” as inspiration. Following are some of opportunities that exist in “Artificial Economics Optimization” field:

- 1) International Association of Artificial Economics Optimization (IAAEO)
- 2) Transactions on Artificial Economics Optimization (TAEO)
- 3) International Journal of Artificial Economics Optimization (IAEO)
- 4) International Conference on Artificial Economics Optimization (ICAEO)
- 5) [www.ArtificialEconomicsOptimization.com](http://www.ArtificialEconomicsOptimization.com)

- 6) B.Tech in Artificial Economics Optimization
- 7) M.Tech in Artificial Economics Optimization
- 8) PhD in Artificial Economics Optimization
- 9) PostDoc in Artificial Economics Optimization
- 10) Artificial Economics Optimization Labs
- 11) To become “Father of Artificial Economics Optimization” field

## 6 Acknowledgements

Thanks to the following members in Editorial Team of TMLAI for accepting and publishing innovative papers in Artificial Human Optimization field:

Professor Farouk Yalaoui	Dr Zdenek Zdrahal
Professor Jai N Singh	Dr Steve S. H. Ling
Dr Kyriakos G Vamvoudakis	Dr Marianne Hoogeveen
Dr Zezhi Chen	Dr Hooshang Hemami
Dr Luis Rodolfo Garcia	Dr Hiroyuki Utsunomiya
Dr Hafiz M. R. Khan	Dr Xiaocong Fan
Dr Nikolaos Georgantas	Dr Chen Yanover
Dr Zeyad Al-Zhour	Dr Vandana Janeja

## REFERENCES

- [1] Satish Gajawada; Entrepreneur: Artificial Human Optimization. Transactions on Machine Learning and Artificial Intelligence, Volume 4 No 6 December (2016); pp: 64-70
- [2] Satish Gajawada, “CEO: Different Reviews on PhD in Artificial Intelligence”, Global Journal of Advanced Research, vol. 1, no.2, pp. 155-158, 2014.
- [3] Satish Gajawada, “POSTDOC : The Human Optimization”, Computer Science & Information Technology (CS & IT), CSCP, pp. 183-187, 2013.
- [4] Satish Gajawada, “Artificial Human Optimization – An Introduction”, Transactions on Machine Learning and Artificial Intelligence, Volume 6, No 2, pp: 1-9, April 2018.
- [5] Liu H, Xu G, Ding GY, Sun YB, “Human behavior-based particle swarm optimization”, The Scientific World Journal, 2014.
- [6] Satish Gajawada, Durga Toshniwal, Nagamma Patil and Kumkum Garg, “Optimal Clustering Method Based on Genetic Algorithm,” International Conference on Soft Computing for Problem Solving (SocPros - 2011), Springer.

- [7] Satish Gajawada, Durga Toshniwal, "Projected Clustering Using Particle Swarm Optimization," International Conference on Computer, Communication, Control and Information Technology (C3IT - 2012), Elsevier.
- [8] Satish Gajawada, Durga Toshniwal, "GAP: Genetic Algorithm Based Projected Clustering Method", 21<sup>st</sup> International Conference on Software Engineering and Data Engineering (SEDE 2012), USA.
- [9] Satish Gajawada, Durga Toshniwal, "Projected Clustering Particle Swarm Optimization and Classification", International Conference on Machine Learning and Computing (ICMLC-2012), Hong Kong.
- [10] Satish Gajawada, Durga Toshniwal, "A Method of Initialization for Genetic Algorithm Based Clustering Technique," International Conference on Computer Science and Information Technology (ICCSIT 2012), interscience, Guwahati.
- [11] Satish Gajawada, Durga Toshniwal, "SPPS: Supervised Projected Clustering Method Based on Particle Swarm Optimization", International Journal of Machine Learning and Computing (IJMLC), vol 2, no 3, 2012.
- [12] Satish Gajawada, Nischey Grover, M.V. Kartikeyan, "Design optimization of non linear tapers for high power gyrotrons using hybrid space mapping techniques", 12<sup>th</sup> IEEE International Vacuum Electronics Conference (IVEC 2011), IEEE.
- [13] Satish Gajawada, Durga Toshniwal, "VINAYAKA: A Semi-Supervised Projected Clustering Method Using Differential Evolution," International Journal of Software Engineering and Applications (IJSEA), 2012.
- [14] Satish Gajawada, Durga Toshniwal, "A framework for classification using genetic algorithm based clustering", The International Conference on Intelligent Systems Design and Applications (ISDA), 2012, IEEE.

# Determinants of Environmental Sanitation in the City of Tshwane, South Africa

Yohannes Worku and Mammo Muchie

*Tshwane University of Technology, Pretoria, South Africa*

[Johannes.Worku@gmail.com](mailto:Johannes.Worku@gmail.com)

## ABSTRACT

The purpose of the study was to identify and quantify key predictors of satisfactory municipal services on municipal waste management in the City of Tshwane, South Africa. The expected and perceived quality of waste management services were analysed by using SERVQUAL analysis. The study found that the perception and expectation of respondents on the quality of sanitary services provided to them were significantly influenced by the degree of motivation of employees of the City of Tshwane at work, the ability of employees of the City of Tshwane to treat all customers with respect, the ability of employees of the City of Tshwane to provide adequate answers promptly to queries raised by customers, and the degree to which employees of the City of Tshwane were skilled on technical issues, in a decreasing order of strength.

**Key words:** Environmental sanitation, City of Tshwane, Service quality, Gap score

## 1 Introduction and background to study

The purpose of study was to assess and evaluate predictors of efficiency in environmental sanitation in the City of Tshwane, South Africa. The general public requires satisfactory service delivery in terms of environmental sanitation and cleanliness. Protests occur over poor sanitary and municipal services from time to time in South African municipalities including the City of Tshwane. The ability to provide satisfactory sanitation services is a key requirement of all employees of the City of Tshwane. Public health and environmental sanitation specialists have shown that developing municipalities experiencing rapid influx from rural areas must have the capacity to enforce environmental and sanitary regulations and bylaws vigorously in order to keep streets clean [1]. The City of Tshwane is characterised by a huge influx of migrants who come to the city in search of jobs and livelihood. The population size of the City of Tshwane has grown significantly since South Africans elected their first democratic government in April 1994. However, the capacity of the City of Tshwane to manage litter and waste properly and ensure environmental sanitation and cleanliness has not improved much in the same period. The City of Tshwane has introduced various tools that could be used for assessing sanitary service quality and customer satisfaction.

Studies conducted by development economists have shown that the task of transforming environmental behaviour among inhabitants of developing cities in Sub-Saharan African countries requires the determination of local municipalities to enforce municipal bylaws with commitment and the clear demonstration of good leadership and management skills [2, 3, 4, 5]. It has also been pointed out that socioeconomic inequalities often undermine cleanliness and environmental sanitation in developing cities

[6, 7]. An understanding of customer opinions relating to efficiency in environmental sanitation and cleanliness are essential for enhancing the quality of sanitary services [8]. Service quality, and the application of efficiently designed services delivery systems, constitutes a penetrating strategy in ensuring cleanliness and environmental sanitation. It is also beneficial for boosting ecotourism [9]. The study aims to assess and evaluate the extent to which inhabitants are satisfied with the quality of services provided by the City of Tshwane. Findings of the study could contribute to efforts made by the City of Tshwane to improve environmental sanitation and cleanliness. The key benefits of maintaining environmental cleanliness and proper waste management in large metropolitan cities such as the City of Tshwane is the ability of local municipalities to attract viable businesses into city centres as a means of growing tax revenue and job creation, and the optimisation of municipal service delivery [10, 11, 12, 13, 14, 15].

According to the City of Tshwane [16], about 1, 734, 295 tons of solid waste is collected each year from businesses operating in the city. The solid waste produced by businesses in the city includes trash or garbage such as wood, product packaging, empty bottles, used tyres and car parts, and cans, garden refuse, furniture, clothing, leftover food, newspapers, wires, grease, appliances, paint, pieces of metal, broken containers, sheet metal and expired medicine. These businesses produce massive volumes of solid and liquid waste on a daily basis. Taxi ranks, bus stations, open flea markets, food outlets, and small businesses located in Pretoria are synonymous with litter, uncontrolled solid and liquid waste, as well as lack of capacity in the efficient management of waste. Passengers travelling in buses and taxis throw out trash through windows of moving vehicles. The collection, disposal and processing of waste produced by businesses and households is regulated by legislative policies. It has been shown that the use of an integrated solid waste management system is essential for reducing the overall cost of waste management in developing cities worldwide [17]. A similar finding is that the promotion of health education on environmental sanitation and primary health care and a strict enforcement of municipal bylaws are vital for ensuring overall environmental cleanliness and the efficient management of waste in developing cities [18]. The study was prompted by a host of factors that are known to undermine environmental sanitation, cleanliness and efficiency in waste management in the City of Tshwane. Examples of such factors are poor infrastructural capacity, poor awareness about the benefits of proper waste management, lack of socioeconomic incentives to stakeholders relevant to the waste management chain, failure to adequately utilize modern waste management and processing technology, failure to vigorously enforce municipal bylaws on sanitation and waste disposal, and the absence of an integrated waste management system in Tshwane [19]. Based on experience learned from members of the European Union, local municipalities must use highly innovative and modern technological methods of maintaining environmental sanitation and cleanliness in urban centres. The authors have argued that the private sector should be given economic incentives to ensure highly efficient and affordable municipal service delivery [20].

### **Objective of study**

The overall objective of study was to assess the degree to which residents of the City of Tshwane are satisfied with the quality of sanitary services that are provided to them by the City of Tshwane.

## 2 Materials and methods of study

Data was collected from a stratified random sample of 1, 012 residents of the City of Tshwane on 22 indicators that are commonly used for the assessment of service quality by the City of Tshwane and other municipalities. Data was collected from each of the 1, 012 residents who were selected for the study by using a structured, pre-tested and validated questionnaire of study consisting of 22 indicators of service quality. The questionnaire of study consisted of 5 dimensions of expectation and perception (reliability, assurance, tangibles, empathy and responsiveness). Each of the respondents in the study had to provide answers to 22 questions related to expectations plus 22 questions related to perceptions. As such, each of the respondents had to provide answers to 44 questions (22 questions on expectation + 22 questions on perception). Measurements of expectations and perceptions were done by using a 5-point ordinal scale. Face validity was used for ensuring validity. The Cronbach Alpha test was used for ensuring reliability and internal consistency. The expected and perceived quality of sanitary services that were provided to the general public by employees of the City of Tshwane was analysed by using SERVQUAL analysis. This was done by estimating gap scores (the average difference between expected and perceived scores).

The 5 dimensions of SERVQUAL analysis are tangibles, reliability, responsiveness, assurance and empathy [21]. The 5 dimensions are defined as follows:

- Tangibles: The appearance of physical facilities, equipment and personnel
- Reliability: The ability to perform the promised service dependably and accurately
- Responsiveness: The willingness to help customers
- Assurance: The knowledge and courtesy of employees and their ability to convey trust and confidence
- Empathy: The provision of caring, individualized attention to customers

Service quality is an achievement in customer service. It reflects at each service encounter. Customers form service expectations from past experiences, word of mouth and advertisement. The perception of residents on the quality of environmental and sanitary services depends on a number of socioeconomic factors [22]. The South African Auditor General (2017) has identified numerous areas of environmental sanitation that need improvement [23].

## 3 Results of study

Table 1 shows frequency proportions that indicate the general characteristics of the participants of study. It can be seen from the table that 854 of the 1, 012 respondents who were selected for the study (84.37%) had a positive overall perception on the quality of services provided to them by employees of the City of Tshwane. Only 158 of the 1, 012 respondents (15.63%) had an overall negative perception. The table shows that 56.62% of the participants of the study were male, whereas the remaining 43.38% were female. The table also shows that 49.01% of the 1, 012 respondents who took part in the study had lived in the City of Tshwane between 11 and 20 years.



**Table 1: General characteristics of respondents (n=1, 012)**

<b>Variable of study</b>	<b>Percentage</b>
Overall perception of residents on the quality of sanitary services	Positive: 854 (84.37%) Negative: 158 (15.63%)
Gender of respondents	Male: 56.62% Female: 43.38%
Age category of respondents	20 or younger: 16.90% 21 to 30: 31.42% 31 to 50: 33.20% 51 or older: 18.48%
Duration of stay in neighbourhood in years	5 or less: 36.17% 6 to 10: 6.62% 11 to 20: 49.01% 21 or more: 8.20%
Highest level of education	Matric level or less: 5.53% Certificate: 27.57% Diploma: 39.23% Bachelor's degree: 15.22% Master's degree or above: 12.45%
Have you ever complained about poor quality sanitary services that you received from the City of Tshwane?	Yes: 4.94% No: 95.06%
Marital status	Single: 28.16% Married: 37.85% Divorced: 30.43% Widowed: 1.48% Others: 2.08%
Job category of respondents	Business owner: 41.11% Government employee: 11.26% Private sector employee: 7.61% Self-employed consultant: 22.73% Unemployed: 8.30% Others: 8.99%

Table 2 shows percentages for the 5 dimensions (reliability, assurance, tangibles, empathy and responsiveness) that were used for the assessment of expectations and perceptions held by residents and ratepayers on the quality of sanitary services provided to the general public by employees of the City of Tshwane.

**Table 2: Expected and perceived scores for dimensions (n=1, 012)**

	<b>Dimension</b>	<b>Expectation score</b>	<b>Perception score</b>
	<b>Reliability: Ability to perform service dependably and accurately</b>		
rel1	Employees of the City of Tshwane are committed for handling all operations in our community	Agree: 97.63% Disagree: 2.37%	Agree: 94.66% Disagree: 5.34%
rel2	Employees are adequately equipped and skilled enough to do their job well	Agree: 99.01% Disagree: 0.99%	Agree: 94.57% Disagree: 5.43%
rel3	Residents who need sanitary services can always depend on employees of the City of Tshwane	Agree: 98.91% Disagree: 1.09%	Agree: 96.74% Disagree: 3.26%
rel4	Employees perform their duty accurately at all times	Agree: 99.01% Disagree: 0.99%	Agree: 96.15% Disagree: 3.85%
	<b>Assurance: Ability of staff to inspire confidence and trust</b>		
ass1	Employees are trustworthy in what they do for the community	Agree: 99.70% Disagree: 0.30%	Agree: 96.34% Disagree: 3.66%
ass2	Employees inspire confidence in what they do for the community	Agree: 99.31% Disagree: 0.69%	Agree: 96.64% Disagree: 3.36%
ass3	Employees are highly punctual and effective at what they do for the community	Agree: 100.00% Disagree: 0.00%	Agree: 96.74% Disagree: 3.26%
ass4	Employees are highly dedicated to the residents who rely on their services	Agree: 100.00% Disagree: 0.00%	Agree: 96.15% Disagree: 3.85%
ass5	Employees are highly professional and disciplined in the course of assisting people in our community	Agree: 99.41% Disagree: 0.59%	Agree: 96.54% Disagree: 3.46%
	<b>Tangibles: Physical facilities, equipment and staff appearance</b>		
tan1	Physical facilities used by employees are appropriate for providing sanitary services	Agree: 100.00% Disagree: 0.00%	Agree: 97.53% Disagree: 2.47%
tan2	Equipment used by employees are appropriate enough for providing adequate sanitary services	Agree: 99.41% Disagree: 0.59%	Agree: 97.33% Disagree: 2.67%
tan3	Employees are physically fit for handling sanitary operations	Agree: 99.70% Disagree: 0.30%	Agree: 99.11% Disagree: 0.89%
tan4	Employees are appropriately trained for handling sanitary services	Agree: 99.70% Disagree: 0.30%	Agree: 96.94% Disagree: 3.06%

	<b>Empathy: The extent to which caring individualized service is given</b>		
emp1	Employees care enough for every customer in our community	Agree: 100.00% Disagree: 0.00%	Agree: 95.65% Disagree: 4.35%
emp2	Employees provide individualized assistance and care to each person who needs their assistance	Agree: 99.31% Disagree: 0.69%	Agree: 96.05% Disagree: 3.95%
emp3	Employees are aware of the needs of all people who require their services	Agree: 100.00% Disagree: 0.00%	Agree: 97.04% Disagree: 2.96%
emp4	Employees do not discriminate based on the personal characteristics of customers	Agree: 100.00% Disagree: 0.00%	Agree: 98.22% Disagree: 1.78%
emp5	Employees show enough appreciation for the pain and suffering endured by customers	Agree: 100.00% Disagree: 0.00%	Agree: 98.32% Disagree: 1.68%
	<b>Responsiveness: Willingness to help and respond to customer need</b>		
res1	Employees are willing to assist all customers in our community at all times	Agree: 100.00% Disagree: 0.00%	Agree: 97.13% Disagree: 2.87%
res2	Employees are prepared to provide assistance to customers under all possible circumstances	Agree: 100.00% Disagree: 0.00%	Agree: 94.86% Disagree: 5.14%
res3	Employees respond to all queries who need their help promptly and with vigour	Agree: 100.00% Disagree: 0.00%	Agree: 96.44% Disagree: 3.56%
res4	The methods and strategies used by employees for providing assistance to customers are appropriate, effective and reliable at all times	Agree: 100.00% Disagree: 0.00%	Agree: 96.34% Disagree: 3.66%

It can be seen from the table that expected and perceived percentage scores were generally similar in magnitude for most of the 5 dimensions. Table 3, below, shows 10 significant two-by-two associations obtained from Pearson's chi-square tests of associations. At the 5% level of significance, significant associations have large observed chi-square values and P-values that are smaller than 0.05. Significant results obtained from Pearson's chi-square tests of associations ( $P < 0.05$ ) showed that overall satisfaction with the quality of sanitary services that were provided by the City of Tshwane was significantly associated with the perception of customers on the following 10 variables of study:

1. Employees respond adequately to all queries promptly and with vigour
2. Employees are adequately equipped and skilled enough to do their job well
3. Employees do not discriminate based on the personal characteristics of people
4. Employees are prepared to provide assistance to all members of the community under all possible circumstances
5. Employees are willing to assist all members of the community at all times

6. Employees are aware of the needs of all people who require their services
7. Employees are highly punctual and effective at what they do for the community
8. Employees provide individualized assistance and care to each person who needs their assistance
9. Employees are aware of the needs of all people who require their services
10. Equipment used by employees are appropriate enough for routine operations

**Table 3: Results obtained from Pearson's chi-square tests of associations**

Overall satisfaction with the quality of sanitary services	Observed Pearson chi-square value	P-value
Employees respond adequately to all queries promptly and with vigour	54.1154	0.0000***
Employees are adequately equipped and skilled enough to do their job well	51.0236	0.0000***
Employees do not discriminate based on the personal characteristics of people	48.2358	0.0000***
Employees are prepared to provide assistance to all members of the community under all possible circumstances	46.2105	0.0000***
Employees are willing to assist all members of the community at all times	41.2233	0.0000***
Employees are aware of the needs of all people who require their services	39.3619	0.0000***
Employees are highly punctual and effective at what they do for the community	36.2546	0.0000***
Employees provide individualized assistance and care to each person who needs their assistance	32.2258	0.0000***
Employees are aware of the needs of all people who require their services	31.2359	0.0000***
Equipment used by employees are appropriate enough for routine operations	30.0123	0.0000***

Legend: Significance levels at \* P<0.05; \*\* P<0.01; \*\*\* P<0.001

It can be seen from Table 3 that all 10 factors are highly significant at the 1% level of significance. This is because all 10 P-values are significantly smaller than 1% = 0.01. The Pearson chi-square test of association is commonly used as a screening tool in cases where the number of variables of study is large. The results obtained above in Table 3 were used for subsequent analysis was done by using factor analysis. As part of

SERVQUAL analysis, comparison was made among paired samples by using the two-sample paired t-test. The comparison made was between the perceptions and expectations of respondents on the quality of sanitary services that were provided to them by employees of the City of Tshwane. All paired t-tests were performed at the 5% level of significance. At the 5% level, true average differences between the two groups being compared with each other were said to be significant if the P-value was less than 0.05. True average differences between the two groups being compared with each other were said to be insignificant if the P-value was greater than or equal to 0.05.

The Cronbach Alpha test [24] was used for ensuring reliability and internal consistency in the measurement tools used for the assessment of expected and perceived values from respondents. Table 4 shows estimated Cronbach Alpha coefficients for expected and perceived values. It can be seen from the table that all estimated coefficients for expected and perceived values by respondents have magnitudes of 75% or above. It can also be seen from the table that estimated coefficients for expected and perceived values were fairly well similar with each other. This shows that the tools used for the assessment of expected and perceived values of the 5 dimensions in the study (reliability, assurance, tangibles, empathy and responsiveness) were fairly highly reliable and suitable for the purpose of the study [25].

**Table 4: Cronbach Alpha coefficients for expected and perceived values**

Dimension	Number of items	Coefficients for expected values	Coefficients for perceived values
Reliability	4	0.8011	0.8109
Assurance	5	0.7845	0.7902
Tangibles	4	0.7759	0.7784
Empathy	5	0.7616	0.7688
Responsiveness	4	0.7584	0.7596

Table 5 shows estimated gap scores for expected and perceived values. A gap score is defined as the difference between the mean of perceived and expected values [25].

Gap score = Perception mean score – Expectation mean score

$$\text{Average gap score} = \frac{\sum_{i=1}^k (\bar{P}_i - \bar{E}_i)}{k} \quad \text{where } k \text{ denotes the number of items used for assessment of}$$

dimensions. In this study, the statistical significance of gap scores was assessed by using P-values obtained from the two-sample paired t-test [24]. At the 5% level of significance, a gap score is said to be statistically significant if the P-value is less than 0.05. If the P-value is greater than or equal to 0.05, a gap score is said to be statistically insignificant. Table 5 shows gap scores estimated from analyses. It can be seen from the table that 20 of the 22 gap scores were significant at the 5% level of significance. There were only 2 items (out of a total of 22 items) that did not produce significant gap scores. These 2 items were items 2 and 3 of the dimension on tangibles. Item 2 of the dimension on tangibles was an assessment on the degree of suitability of the equipment used by employees of the City of Tshwane. Item 3 of the dimension on tangibles was an assessment on the degree of physical fitness of employees of the City of Tshwane for carrying out sanitary services effectively. With the expectation of the 2 gap scores corresponding to these 2 items, all other gap scores (20 out of 22) were statistically significant at the 5% level of significance. At the 5% level of significance, the results show a significant disparity between expected and perceived

values. As such, the City of Tshwane must make the initiative to let stakeholders know more about the sanitary services that are routinely provided to the community as a means of increasing awareness and appreciation about the quality of services provided, and to enable members of the community to utilize the services optimally [25].

**Table 5: Estimated gap scores for expected and perceived values**

DIMENSION	Item	EXPECTATION MEAN SCORE	PERCEPTION MEAN SCORE	GAP SCORE (P – E)	P-value
Reliability	1	1.378446	1.486729	-0.108	0.0000
	2	1.302558	1.504429	-0.202	0.0000
	3	1.296529	1.423322	-0.127	0.0000
	4	1.281149	1.412654	-0.132	0.0000
<b>Total</b>		<b>5.258682</b>	<b>5.827134</b>	<b>-0.568</b>	
<b>Average gap score [Total of (P – E) / 4]</b>				<b>0.14211</b>	
Assurance	1	1.227654	1.325523	-0.098	0.0000
	2	1.201155	1.341121	-0.140	0.0000
	3	1.192245	1.323426	-0.131	0.0000
	4	1.230524	1.314987	-0.084	0.0000
	5	1.202646	1.392145	-0.189	0.0000
<b>Total</b>		<b>6.054224</b>	<b>6.697202</b>	<b>-0.642</b>	
<b>Average gap score [Total of (P – E) / 5]</b>				<b>-0.1284</b>	
Tangibles	1	1.240012	1.400567	-0.161	0.0000
	2	1.271359	1.363521	-0.092	0.0000
	3	1.281456	1.308090	-0.027	0.0001
	4	1.285677	1.387540	-0.102	0.0001
<b>Total</b>		<b>5.078504</b>	<b>5.459718</b>	<b>-0.382</b>	
<b>Average gap score [Total of (P – E) / 4]</b>				<b>0.09550</b>	
Empathy	1	1.284417	1.420608	-0.136	0.0001
	2	1.215674	1.352642	-0.137	0.0000
	3	1.194687	1.431257	-0.237	0.0000
	4	1.220048	1.419080	-0.199	0.0000
	5	1.192671	1.362358	-0.170	0.0000
<b>Total</b>		<b>6.106945</b>	<b>6.985945</b>	<b>-0.879</b>	
<b>Average gap score [Total of (P – E) / 5]</b>				<b>-0.1758</b>	
Responsiveness	1	1.281456	1.482564	-0.201	0.0001
	2	1.262328	1.362546	-0.100	0.0001
	3	1.251169	1.435687	-0.185	0.0000
	4	1.291567	1.552648	-0.261	0.0001
<b>Total</b>		<b>5.086445</b>	<b>5.833445</b>	<b>-0.747</b>	
<b>Average gap score [Total of (P – E) / 4]</b>				<b>-0.1868</b>	

Table 6 compares male and female respondents with regards to the 5 dimensions (reliability, assurance, tangibles, empathy, and responsiveness) based on group mean scores and P-values obtained from two-sample t-tests.



**Table 6: Comparison of dimensions with regards to gender based on t-tests**

Dimension	Mean score for male respondents	Mean score for female respondents	P-value
Reliability	0.229	0.212	0.4609
Assurance	0.163	0.287	0.0083**
Tangibles	0.104	0.096	0.8541
Empathy	0.229	0.212	0.6779
Responsiveness	0.196	0.218	0.5510

Legend: Significance levels at \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$

It can be seen from the table that male and female respondents differed significantly at the 1% level of significance with regards to assurance values. Table 7 compares 4 age categories of respondents with regards to the 5 dimensions (reliability, assurance, tangibles, empathy, and responsiveness) based on group mean scores and P-values obtained from the one-way Analysis of Variance (ANOVA) test. The one-way ANOVA test is appropriate for comparison as the number of age categories is more than 2, and the variables of comparison are continuous [24].

**Table 7: Comparison of dimensions with regards to age category**

Dimension	Mean score for respondents 20 years old or younger	Mean score for respondents 21 to 30 years of age	Mean score for respondents 31 to 50 years of age	Mean score for respondents 51 years of age or older	P-value
Reliability	0.112	0.124	0.279	0.212	0.0149*
Assurance	0.107	0.180	0.281	0.260	0.0566
Tangibles	0.004	0.078	0.164	0.120	0.0517
Empathy	0.150	0.228	0.235	0.250	0.4435
Responsiveness	0.088	0.306	0.211	0.127	0.0001***

Legend: Significance levels at \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$

It can be seen from the gap scores of reliability and responsiveness differed significantly by age category at the 1% level of significance. The difference with regards to reliability gap scores is attributed to differences between age categories (21 to 30) and (31 to 40) based on a P-value of  $0.028 < 0.05$  obtained from Bonferroni's test (Hair, Black, Babin & Anderson, 2010). The difference with regards to responsiveness gap scores is attributed to differences between age categories (20 or less) and (21 to 30) based on a P-value of  $0.000 < 0.05$  as well as age categories (21 to 30) and (51 or above) based on a P-value of  $0.003 < 0.05$  obtained from Bonferroni's test. Factor analysis (Field, 2013: 134-158) was used for identifying influential predictors of perception and expectation. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was used in order to test the adequacy of the sample used for factor analysis, and the test gave an estimated KMO value of  $0.849 = 84.9\%$ , a figure that is greater than 75%. This large figure indicates that results estimated from factor analysis for perception are fairly well reliable. Bartlett's test of Sphericity was used for testing the adequacy of the correlation matrix, and gave an observed chi-squared value of 1046.777 (very large value) with 229 degrees of freedom (very large degrees of freedom) and a P-value of 0.000 (a P-value that is much smaller than 0.05). These estimated figures show that the use of factor analysis for identifying key predictors of perception is fairly well justified and appropriate.

**Table 8: Estimates obtained from the KMO and Bartlett's test for perception**

Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy	0.849
Observed value of chi-square statistic for KMO test	1046.777
Bartlett's Test of sphericity Degrees of freedom	229
P-value for Bartlett's Test of sphericity Degrees of freedom	0.000

The principal axis factoring method was used for estimating communalities for 4 influential predictors of perception. Table 9 shows the communalities estimated for the 4 influential predictor variables of perception.

**Table 9: Communalities extracted for 4 influential predictors of perception**

Variable of study	Extraction based on principal component analysis
Employees do not discriminate based on the personal characteristics of customers	0.701
Employees are aware of the needs of all people who require their services	0.684
Employees are highly punctual and effective at what they do for the community	0.574
Employees provide individualized assistance and care to each person who needs their assistance	0.536

Table 10 shows estimated Eigen values [26] and percentages of explained variation for the 4 key predictors of perception. Based on results obtained from factor analysis for expectations, the expectation of respondents on the quality of sanitary services that were provided to them was significantly influenced by 4 key predictors of perception. These 4 predictor variables were the ability of employees not to discriminate among customers based on personal characteristics, the ability of employees to be aware of the needs of all people who require their services, the ability of employees to be punctual and effective, and the ability of employees to provide individualized assistance and care to each person who needs their assistance, in a decreasing order of strength. It can be seen from the table that the cumulative variation explained by the 4 influential variables is equal to 77.448%, a figure which is larger than 75%. This indicates that the 4 extracted factors account for variability in perception adequately enough.

**Table 10: Eigen values estimated from factor analysis for perception**

Variable	Eigen value	Percentage of explained variance	Cumulative percentage of explained variance
Employees do not discriminate based on the personal characteristics of customers	2.889	30.119	30.119
Employees are aware of the needs of all people who require their services	2.809	15.884	46.003
Employees are highly punctual and effective at what they do for the community	2.771	15.449	61.452
Employees provide individualized assistance and care to each person who needs their assistance	2.003	15.996	77.448

#### 4 Discussion of results

The study has found that 84.37% of the 1, 012 respondents who took part in the study were satisfied with the overall quality of sanitary services that were provided to them by the City of Tshwane. Only 15.63% of respondents were not satisfied with the overall quality of services provided to them. The study showed that most of the respondents had a positive perception on the quality of routine sanitary services such as water and lights and waste removal by employees of the City of Tshwane. The study found that as many as 87.13% of respondents had a positive perception about the degree of commitment shown to them by employees of the City of Tshwane. Based on results obtained from SERVQUAL analysis, 20 of the 22 gap scores were found to be significant at the 5% level of significance. There were only 2 items (out of a total of 22 items) that did not produce significant gap scores. These 2 items were items 2 and 3 of the dimension on responsiveness. Item 2 of the dimension on responsiveness was an assessment on the degree of suitability of the equipment used by sanitary employees for carrying out routine services. Item 3 of the dimension on responsiveness was an assessment on the degree of physical fitness of employees of the City of Tshwane for carrying out routine sanitary services effectively. With the expectation of the 2 gap scores corresponding to these 2 items, all other gap scores (20 out of 22) were statistically significant at the 5% level of significance.

Based on results obtained from factor analysis, the perception and expectation of respondents on the quality of sanitary services that were provided to them were significantly influenced by 4 key predictors of perception. These 4 predictor variables were the degree of motivation of employees of the City of Tshwane at work, the ability of employees of the City of Tshwane to treat all customers with respect, the ability of employees of the City of Tshwane to provide adequate answers promptly to queries raised by customers, and the degree to which employees of the City of Tshwane were skilled on technical issues, in a decreasing order of strength. Similar results were obtained from logit analysis. The results showed that the perception and expectation of respondents were influenced by similar variables of study. The key finding of the study is that 854 of the 1, 012 respondents who were selected for the study (84.37%) had a positive overall perception on the quality of sanitary services provided to them by employees of the City of Tshwane. Only 158 of the 1, 012 respondents (15.63%) had an overall negative perception. The study

also showed that 92.66% of respondents had a positive perception about the degree to which employees of the City of Tshwane were appropriately qualified for the work they had to do for the community. The percentage of respondents who had a negative expectation was only 7.34%. The study showed that 95.55% of respondents had a positive perception about the degree to which employees of the City of Tshwane provide individualized assistance and care to people who need their assistance. The percentage of respondents who had a negative expectation was only 4.45%. Based on results obtained from the two-sample paired t-test, 20 of the 22 gap scores used in the study were significant at the 5% level of significance. There were only 2 items (out of a total of 22 items) that did not produce significant gap scores. These 2 items were items 2 and 3 of the dimension on tangibles. Item 2 of the dimension on tangibles was an assessment on the degree of suitability of the equipment used by employees of the City of Tshwane. Item 3 of the dimension on tangibles was an assessment on the degree of physical fitness of employees of the City of Tshwane for carrying out sanitary services effectively. With the expectation of the 2 gap scores corresponding to these 2 items, all other gap scores (20 out of 22) were statistically significant at the 5% level of significance. At the 5% level of significance, the results show a significant disparity between expected and perceived values [25]. As such, the City of Tshwane must make the initiative to let stakeholders know more about the sanitary services that are routinely provided to the community as a means of increasing awareness and appreciation about the quality of services provided, and to enable members of the community to utilize the services optimally.

Gap scores estimated from the paired t-test showed that the gap scores of reliability and responsiveness differed significantly by age category at the 1% level of significance. The difference with regards to reliability gap scores is attributed to differences between age categories (21 to 30) and (31 to 40) based on a P-value of  $0.028 < 0.05$  obtained from Bonferroni's test [26]. The difference with regards to responsiveness gap scores is attributed to differences between age categories (20 or less) and (21 to 30) based on a P-value of  $0.000 < 0.05$  as well as age categories (21 to 30) and (51 or above) based on a P-value of  $0.003 < 0.05$  obtained from Bonferroni's test. Results obtained from factor analysis for expectations showed that the expectations of respondents on the quality of sanitary services that were provided to them were significantly influenced by 4 key predictors of perception. These 4 predictor variables were the ability of employees not to discriminate among customers based on personal characteristics, the ability of employees to be aware of the needs of all people who require their services, the ability of employees to be punctual and effective, and the ability of employees to provide individualized assistance and care to each person who needs their assistance, in a decreasing order of strength. Similar findings were obtained from binary logistic regression analysis [27]. Studies conducted by the World Bank [28], Buckley and Ghauri [29], Alexander [30], Akinboade, Mokwena and Kinack [31], [32], [33], [34], [35], [36], [37], [38], [39], [40] have shown that failure to provide satisfactory sanitary and waste management services to ratepayers, inhabitants and entrepreneurs living and conducting business in major metropolitan centres such as the City of Tshwane often leads to municipal protests over poor service delivery, and failure to attract viable business enterprises into South Africa. King (2014) has published a code of good corporate governance and leadership that should be followed by all municipal officials, project leaders and civil servants working in the 205 local municipalities in South Africa [41, 42].

Results obtained from factor analysis for perceptions showed that the perceptions of respondents on the quality of sanitary services that were provided to them were significantly influenced by 4 key predictors of perception. These 4 predictor variables were the ability of employees not to discriminate among residents based on the personal characteristics of customers, the ability of employees to be aware of the needs of all people who require their services, the ability of employees to be punctual and effective at what they do for the community, and the ability of employees to provide individualized assistance and care to each person who needs their assistance, in a decreasing order of strength. Results obtained from logit analysis showed that the degree of satisfaction of customers with the quality of sanitary services provided to them was significantly influenced by 4 predictor variables. These predictor variables were the degree of motivation of employees of the City of Tshwane at work, the ability of employees of the City of Tshwane to treat all customers with respect, the ability of employees of the City of Tshwane to provide adequate answers promptly to queries raised by customers, and the degree to which employees of the City of Tshwane were skilled on technical aspects of their job.

## REFERENCES

- [1] Aguilar, O. M., & Krasny, M. E. (2011). Using the Communities of Practice Framework to Examine an after-School Environmental Education Program for Hispanic Youth. *Environmental Education Research*, 17(2), 217–233.
- [2] Brown, B. A. (2011). Isn't That Just Good Teaching? Disaggregate Instruction and the Language Identity Dilemma." *Journal of Science Teacher Education*, 22(8), 679–704.
- [3] Carmi, N., Arnon, S., & Orion, N. (2015). Transforming Environmental Knowledge into Behavior: The Mediating Role of Environmental Emotions. *The Journal of Environmental Education*, 46(3), 183–201.
- [4] Fauconnier, A., & Mathur-Helm, B. (2013). Black economic empowerment in the South African industry: A case study of Exxaro limited. *South African Journal of Business Management*, 39(4), 1-14.
- [5] Elliot, R., & Boshoff, C. (2013). The influence of organisation factors in small tourism businesses on the success of internet marketing. *Management Dynamics*, 14(3), 44 -58.
- [6] Fletcher, R. (2012). Using the Master's Tools? Neoliberal Conservation and the Evasion of Inequality. *Development and Change*, 43(1), 295–317.
- [7] Forestelli, R. M., & Angulo-Aguilar, J. E. (2013). *Panorama socioeconómico de los cantones de Osa & Golfito: tendencias y desafíos para el desarrollo sostenible*. Stanford, CA: INOGO, Stanford Woods Institute for the Environment.

- [8] Gruenewald, D. A., & Smith, G. A. (2014). *Place-based Education in the Global Age: Local Diversity*. New York: Routledge.
- [9] Hunt, C. A., Durham, W. H., Driscoll, L., & Honey, M. (2015). Can Ecotourism Deliver Real Economic, Social, and Environmental Benefits? A Study of the Osa Peninsula, Costa Rica. *Journal of Sustainable Tourism*, 23(3), 339–357.
- [10] Grant, R. M. (2013). The Resource-based Theory of Competitive Advantage. *California Management Review*, 33(3), 114-135.
- [11] Bloodgood, J. M. (2012). Organizational routine breach response and knowledge management. *Business Process Management Journal*, 18(3), 376-399.
- [12] Boyne, G. J. O., John, P., & Petrovsky, N. (2010). Does public service performance affect top management turnover? *J Public Adm Res Theory*, 20(2), 261-279.
- [13] Hood, C., & Dixon, R. (2010). The political payoff from performance target systems: no-brainer or no-gainer? *J Public Adm Res Theory*, 20(2), 281-298.
- [14] Bhatt, G. D., & Grover, V. (2013). Types of Information Technology Capabilities and Their Role in Competitive Advantage: An Empirical Study. *Journal of Management Information Systems*, 22(2), 253-277.
- [15] Globerman, S., Mike, W. P., & Daniel, M. S. (2011). Corporate governance and Asian companies. *Asia Pacific Journal of Management*, 28(1), 1-14.
- [16] City of Tshwane. 2017. *Annual report on the City of Tshwane for 2015/2016*. Pretoria: City of Tshwane. Retrieved from <http://www.tshwane.gov.za/AboutTshwane/>
- [17] Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., Narayan, R., & Law, K. L. (2015). Plastic waste inputs from land into the ocean. *Science*, 347(6223), 768-771.
- [18] Lohri, C.R., Camenzind, E. R., & Zurbrugg, C. (2014). Financial sustainability in municipal solid waste management—Costs and revenues in Bahir Dar, Ethiopia. *Waste Management*, 34(2), 542-552.
- [19] Worku, Y., Snyman, J., & Muchie, M. (2014). The management of municipal solid waste generated by businesses operating in the city of Tshwane, South Africa. *European Journal of Engineering and Technology*, 2(2), 77-106.



- [20] Karlstrom, H., & Finstad, T. (2013). The Stedegenhet of Nordic STS. *Nordic Journal of Science and Technology*, 1(1), 5-11.
- [21] Zeithaml, V., Bitner, M. J., & Gremler, D. (2009). *Service Marketing: Integrating Customer Focus Across the firm*. 5th ed. Boston: McGraw International.
- [22] Meeusen, R. (2014). Exercise, nutrition and the brain. *Sports Medicine*, 44(1), 47-56.
- [23] South African Auditor-General. (2017). *Audit report on Tshwane municipal service delivery for the financial year 2015/2016*. Retrieved from <http://hsf.org.za/resource-centre/> Pretoria: South African Auditor-General.
- [24] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective*. London: Pearson.
- [25] Parasuraman, A. (1998). Customer service in business-to-business market. *Journal of Business and Industrial Marketing*, 13(1), 310-321.
- [26] Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. 3rd ed. London: SAGE Publications.
- [27] Hosmer, D. W., & Lemeshow, S. (2013). *Applied Logistic Regression Analysis*. New York: John Wiley & Sons.
- [28] World Bank. (2015). *World development report: making services work for poor people*. Washington DC: World Bank.
- [29] Buckley, P. J., & Ghauri, P. N. (2012). Globalisation, Economic Geography and the Strategy of Multinational Enterprises. *Journal of International Business Studies*, 35(2), 81-98.
- [30] Alexander, P. (2010). Rebellion of the poor: South Africa's service delivery protests—a preliminary analysis. *Review of African Political Economy*, 37(123), 25-40.
- [31] Akinboade, O. A., Mokwena, M. P. & Kinfaek, E. C. (2013). Understanding citizens' participation in service delivery protest in South Africa. *International Journal of Social Economics*, 40(5), 461-462.
- [32] Khale, S., & Worku, Z. (2013). Factors that affect municipal service delivery in Gauteng and North West Provinces of South Africa. *African Journal of Science, Technology, Innovation and Development*, 2(3), 61-70.

- [33] Khale, S., & Worku, Z. (2015). Benefits of good corporate governance principles: A study of the City of Tshwane, South Africa. *Journal of Corporate Governance and Control*, 13(1), 753-770.
- [34] Worku, Z. (2015). Barriers to the development of entrepreneurial activities in newly established small, micro and medium-sized enterprises in Tshwane. *Journal of Corporate Ownership and Control*, 3(12), 332-340.
- [35] Municipal IQ Hotspot Monitor. (2017). *Press Release: Municipal IQ's Municipal Hotspots results for 2016*. Pretoria: Municipal IQ Hotspot Monitor. Retrieved from <http://www.municipaliq.co.za/>
- [36] Statistics South Africa. (2017). *Fast facts and figures*. Retrieved from [www.statssa.gov.za](http://www.statssa.gov.za) Pretoria: Statistics South Africa.
- [37] Kovacevic, N. (2012). Righting wrongs: Affirmative action in South Africa. *Harvard International Review*, 6(1), 12-15.
- [38] Dzansi, D. Y., Dzansi, L. W. (2010). Understanding the impact of human resource management practices on municipal service delivery in South Africa: An organizational justice approach. *African Journal of Business Management*, 4(6), 995- 1005.
- [39] Fahlenbrach, R., & Stulz, R. M. (2011). Bank CEO incentives and the credit crisis. *Journal of Financial Economics*, 99(1), 11–26.
- [40] City of Cape Town. (2017). *Annual report on the City of Cape Town for 2015/2016*. Cape Town: City of Cape Town. Retrieved from <http://www.capetown.gov.za/en/reports/Pages/default.aspx>
- [41] Khale, S. (2015). Assessment of the quality of municipal services in the City of Tshwane, South Africa. *Journal of Corporate Governance and Control*, 13(1), 678- 695.
- [42] King, M. E. (2014). *The King report on corporate governance for South Africa*. Johannesburg: Institute of Directors in Southern Africa.

# Classification of Encrypted Texts using Deep Learning

<sup>1</sup>Zeinab Nazemi Absardi, <sup>2</sup>Reza Javidan

<sup>1,2</sup>Department of Computer Engineering and IT, Shiraz University of Technology, Shiraz, Iran;  
zeinab.nazemi@gmail.com; reza.javidan@gmail.com

## ABSTRACT

The most widely used cryptographic systems can identify cryptographic algorithms and identify encryption keys. Statistical methods and learning a variety of machines have been used to identify cryptographic algorithms, each of which has its own advantages and disadvantages. This paper seeks to provide a method for identifying the algorithm used for encrypted texts in text files. Since the volume of this kind of data is very big and increases at any given moment, then the accuracy is calculated by voting of these classifiers. The process of identifying the encryption algorithm is also known from the encrypted texts as the classification of text. So, three methods of encryption AES, RC5, BLOWFISH have been used to evaluate system performance. A three class's classifier is needed, for this purpose, k-nearest neighbor's algorithm has been used. This article is based on a deep learning approach, provides a new method for identifying the pattern in cryptographic texts and learning them by methods of representing features. The proposed method, consists of four parts of the preprocessing, feature learning, data classification and voting. The proposed system's efficiency in algorithm classification is 99.1%.

**Keywords:** Text classification; Encrypted texts; Deep learning; Encryption algorithms.

## 1 Introduction

Due to the growing popularity and increased interaction of social media users through conversational texts, the clustering of these types of texts has become an important task. In this paper, a method for clustering texts through deep learning is proposed.

The most widely used cryptographic systems can identify cryptographic algorithms and identify encryption keys. Statistical and machine learning methods have been used to identify the cryptographic algorithm, each of which has its own advantages and disadvantages. Machine learning is a general-purpose method that can learn relationships from the data without the need to define them a priori [1]. For decades, constructing a machine learning system required careful engineering and domain expertise to transform the raw data into a suitable internal representation from which the learning subsystem, often a classifier, could detect patterns in the data set. Conventional techniques are composed of a single, often linear, transformation of the input space and are limited in their ability to process natural data in their raw form [2]. Deep learning allows computational models that are composed of multiple processing layers based on neural networks to learn representations of data with multiple levels of abstraction [3].

The key aspect of deep learning is that these layers of features are not designed by human engineers, but they are learned from data using a general purpose learning procedure. By using a recently proposed leveled homomorphic encryption scheme, it is possible to delegate the execution of a machine learning algorithm to a computing service while retaining confidentiality of the training and test data. Since the

computational complexity of the homomorphic encryption scheme depends primarily on the number of levels of multiplications to be carried out on the encrypted data, we define a new class of machine learning algorithms in which the algorithm's predictions, viewed as functions of the input data, can be expressed as polynomials of bounded degree [4].

There have been numerous studies on classification of encrypted texts which have led to many different approaches. Most of these approaches use predefined features extracted by an expert in order to classify encrypted texts. In contrast, in this study, a deep learning based approach is proposed which integrates both feature extraction and classification phases into one system.

This article seeks to provide a new way of identifying patterns in cryptographic texts and learning them by way of representing features. The system designed for the three encryption algorithms named AES, RC5 and BLOWFISH has been evaluated. It can also be used for other cryptographic algorithms. Section 2 gives a brief overview of the literature in this area. The third Section deals with the proposed methodology and examines the architecture of the system. The tests carried out and the results obtained are also discussed in section four. Finally, Section Five is dedicated to providing conclusions.

## 2 Review of related works

Classification of encrypted texts has become significantly important with rapid growth of current Internet network and online applications. There have been numerous studies on this topic which have led to many different approaches. In the area of identifying the text encryption algorithm, there are several sections in which the difference in how they implement them has caused differences in the work of various researchers, one of these sections is the choice of how to input the system. Given that the classification of codes based on texts is encrypted. The researchers have used various methods for converting texts into acceptable inputs for the system [5], including the vectorization of strings from texts.

Also, the playback of feature selection and classification is also one of the most influential parts in the design of the classification system of the cryptographic algorithm that results in different results from the performance of the system in different work [6].

One of the most popular classifiers in this area is Naive Bayesian, Support Vector Machine, Neural Network, and Instance Based Learner. The Naive Bayesian classification is one of the heaviest Bayesian classifiers in terms of calculation, but it has two points compared to others, which is the first to be easy to build. Second, the classification process is very efficient [7].

Also, in the area of detecting the encryption algorithm used in the texts, SVM classifier have been used extensively [8].

Neural network classifiers are also very popular in a variety of domains. In the area of detecting encrypted texts, systems based on the classification of the neural network have shown excellent results that add to the acceptability of this type of classification [9-11].

Network traffic classification has become significantly important with rapid growth of current Internet network and online applications. There have been numerous studies on this topic which have led to many different approaches. Most of these approaches use predefined features extracted by an expert in order to classify network traffic [12].

In [14] a new technique developed to provide solutions for running deep neural networks over encrypted data. They develop new techniques to adopt deep neural networks within the practical limitation of current homomorphic encryption schemes. Test results validate the soundness of approach with several convolutional neural networks with varying number of layers and structures. When applied to the MNIST optical character recognition tasks, this approach achieves 99.52% accuracy which significantly outperforms the state-of-the-art solutions and is very close to the accuracy of the best non-private version, 99.77%. The approach also applied to CIFAR-10, which is much more complex compared to MNIST, and were able to achieve 91.5% accuracy with approximation polynomials used as activation functions. These results show that CryptoDL provides efficient, accurate and scalable privacy-preserving predictions.

In [13] a novel scheme for a classifier owner is proposed to delegate a remote server to provide the privacy-preserving classification service for users. In the proposed scheme, efficient classification protocols for two concrete classifiers respectively are designed. The prototype of the scheme and conduct tests is implemented. The test results show that the scheme is practical.

### 3 Proposed Method

One of the most important steps in designing classification systems and clustering information is to select a suitable method for extracting and representing features in the raw input data. The introduction of a learning-based approach to non-monitoring features improves system performance and automatically learns the features of input data. The proposed method consists of four parts of the preprocessing, learning features (representing feature and information retrieval based on the characteristics learned), classification of represented data, and voting. The architecture of proposed system is shown in Figure 1.

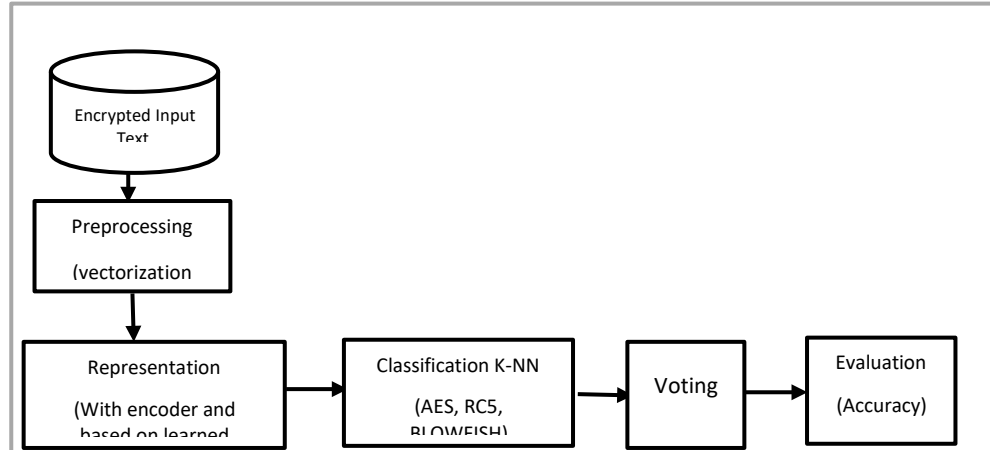


Figure 1. Architecture of proposed method.

#### 3.1 Preprocessing

The preprocessing must first deepen the encrypted text files to match the acceptable input for the neural networks. For this purpose, text vectoring is used. The data used for this study includes three encryption algorithms called AES, RC5 and BLOWFISH, initially similar to the bag of word method in the processing of texts, 500000 strings of 128 bits of the texts of each algorithm are selected. Then put the inputs together and create a collection of 1,500,000 encrypted text strings. The proposed system has two main stages of training and application. At the training stage, the system uses learning resources to learn the best

features available in the database. The architecture of the proposed system in the training phase is shown in Figure 2.

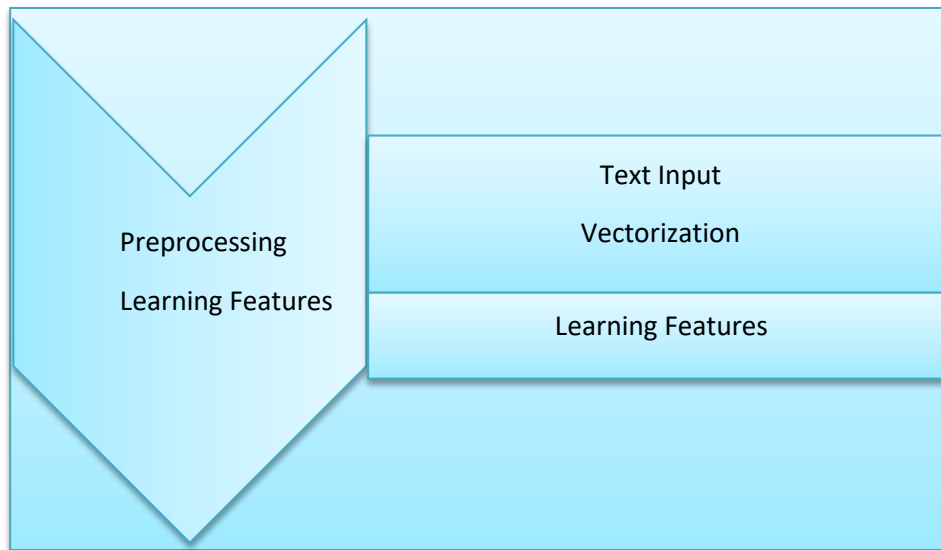


Figure 2. Architecture of the training step.

After learning the features by the cryptography itself, the system is ready to use. In fact, after training, the encoder in the training stage of the system should be able to represent raw data in a new space based on the features learned.

### 3.2 Learning Features

Learning features are the methods used to representing raw data in the form of a new representational model in the new space. Depending on the issue, this new space can be smaller or larger than the original space. To learn features, a shorter linear encoder has been used shown in Figure 3.

This encoder acts on the basis of a gradient reduction, aiming to create a thumbnail representation of the data. The objective function defined for the predecessor and its related definitions is expressed in Formula 1.

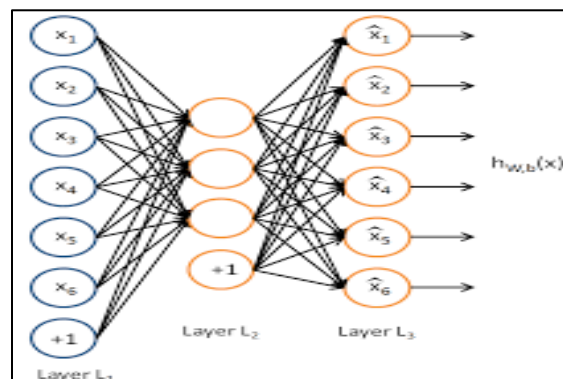


Figure 3. Architecture of encoder in middle layer.

This encoder acts on the basis of a gradient reduction, aiming to create a thumbnail representation of the data. The objective function defined for the predecessor and its related definitions is expressed in Formula 1.

$$\begin{aligned}
 J_{space}(W, b) &= J(W, b) + \beta \sum_{j=1}^{S^*} KL(p \parallel p) \\
 KL(p \parallel \hat{p}_j) &= p \log \frac{a}{\hat{p}_j} + (1 - p) \log \frac{1 - a}{1 - \hat{p}_j} \\
 J(W, b) &= \left[ \frac{1}{m} \sum_{i=1}^m (W, b; x^{(i)}, y^{(i)}) \right] + \frac{\lambda}{2} \sum_{j=1}^{n+1} \sum_{i=1}^n \sum_{j=1}^{n+1} (W_{ji}^{(i)})^2
 \end{aligned} \tag{1}$$

In fact, encoder after training on instructional data, the encoder itself becomes a tool to unravel the raw data in an appropriate space for classification. In other words, encoders are a good alternative to traditional method of feature extraction. Especially in cases such as the cryptographic field where the type and gender of the input data are ambiguous, encoders themselves have a high ability. To implement the proposed system based on deep learning, the encoder itself in Figure 3 is designed in three layers, which includes a coding layer, an intermediate layer, and a decoding layer. The given data in accordance with the preprocessing referred to above has become an acceptable input for the system and is sent to the encoder itself with the architecture shown in Figure 4 to be trained with the activation function RELU and the number of replicates of the network is 100,000.

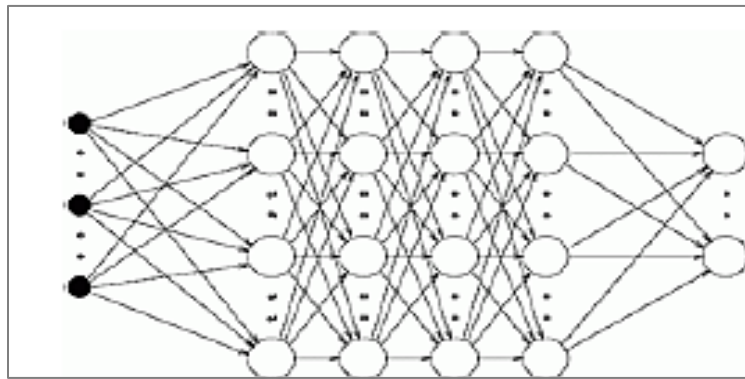


Figure 4. Proposed neural network architecture.

Tag = 3

Decoding layer = 128-256

Intermediate layer = 512

Encoder layer = 128

After learning this encoder, it can be used to represent the raw data in the new space.

### 3.3 Classification

After representing raw inputs, it is time to classify them. In this system, this study is trying to identify the method of encryption of input data. As mentioned above, in this research, RC5 AES and BLOWFISH encryption methods have been used to evaluate system performance. So in this system a class of three classes is needed. For this purpose, a KNN classifier has been used based on the Mahalanobis distance.



Because of its high codec capability in representing raw data, a simple classifier of KNNs has the ability to properly handle data. Also, due to the lack of need for this class of classifier, learning and training, if the number of classifier is changed (changing the number of cryptographic algorithms), this classifier of categories easily adapts itself.

### 3.4 Voting

As stated, the encoder represents the extracted vectors from raw input data using the learned features. After representing these vectors, they classify their categories and determine which algorithms are encrypted. But the purpose of this system is to detect the encryption algorithm of a text. According to the description given, a text contains much number of words that are made in the form of a vector in this system. As a result, to decide on a text, one can extract the results of multiple vector classifications and decide on the results of their classification. This final decision is taken by a majority vote.

## 4 Evaluation of Proposed Method

To evaluate the proposed system, 128-bit characters are randomly generated to be used in the training and testing of the system. Also, to test the system more precisely, two test models were performed on this system. In the first approach, despite the difference in the training and testing data classification, the key to their encryption is the same. In the second method, the cryptographic key of the training data is completely different with the test data. It is important to note that in all of the above methods, the encryption key for the encoder's own learning resources, which is used to extract the property; it is different from other educational and test tutorials in the upper hand. To evaluate the proposed system, 500000 vectors of 128 selected from each algorithm are put together to create a collection of 1,500,000 vectors of encrypted texts by the three algorithms.

a KNN classifier based on  $k = 10$  was used for evaluating the proposed architecture of encoder that the percentage of system performance in the same way as the cryptographic key of the training and test data of the cluster was 84.3%, and in the case of different cryptographic keys, the training and test data of the cluster was 39.9%. Given that at the time of the test, in real terms, every text file received a large number of test inputs. Finally, the final view of the system for that text file is presented, a test was presented based on this. In this test, for each text file, 100 samples of 128 bits are tested, then, by voting in the votes obtained from the system for 100 samples, the final commentary on the input text file is presented. According to the conditions mentioned above, the system performance was equal to 99.1 percent for the same key and 62.5 percent for the encryption key. In Table 1 and 2, the efficiency of the system has been investigated.

**Table 1. Accuracy of classification of encoded vectors.**

Algorithm Key	Accuracy of Classification
same for train and test data	84.3
different for train and test data	39.9

**Table 2. Accuracy of classification of encoded vectors voting.**

Algorithm Key	Accuracy of Classification
same for train and test data	99.1
different for train and test data	62.5

Considering the different conditions of training and testing in the systematic analysis systems, in these studies, there is generally no comparison with past work to implement the proposed system, the CAFFE profound learning framework has been used. Testing and analyzing outcomes have also been done in MATLAB software.

## 5 Conclusion

In this article a new method for classifying encrypted texts is presented. In this way, a deep network was provided to receive input data from various algorithms and to train their features and eventually represent them in a suitable environment. This paper seeks to provide a method for identifying the algorithm used for encrypted texts in text files. Since the volume of this kind of data is very big and increases at any given moment, then the accuracy is calculated by voting of these classifiers. So, three methods of encryption AES, RC5, BLOWFISH have been used to evaluate system performance. Also there are several methods for classifying cryptographic algorithms. In this system a three class's classifier is needed. For this purpose, k-nearest neighbor's algorithm (K-NN) has been used. This article is looking for a deep learning-based approach. In order to analyze the input data, the proposed method consists of four parts of the preprocessing, feature learning, data classification and voting. This paper, seeks to provide a new method for identifying the pattern in cryptographic texts and learning them by methods of representing features. The system designed for the three encryption algorithms named BLOWFISH, RC5, and AES is evaluated and can be generalized for use with other cryptographic algorithms. The proposed system's efficiency in algorithm classification is 98.9%. Subsequently, by the same network trained, the data was represented in a new space and classified by the KNN category and the maximum voting method. Given that this method has been tested for a limited number of cryptographic algorithms, it can be evaluated for a range of algorithms and, if possible, create similar conditions, the results will be compared with other similar work carried out in this area. There are also various parameters in the deep design grid that can be modified and reviewed by their results to optimize the system.

## REFERENCES

- [1] Murphy Kevin P. *Machine learning: a probabilistic perspective*, MIT press, 2012. p. 119-121.
- [2] Bengio Y. et al, *Representation learning: a review and new perspectives*. IEEE Trans Pattern Anal Mach Intell, 2013. 35(4): p. 1798–828,.
- [3] LeCun Y. et al, *Deep learning*, Nature, 2015. p. 436–44.
- [4] Graepel T. et al , *ML Confidential: Machine Learning on Encrypted Data*. Information Security and Cryptology ICISC, 2012.

- [5] Liwen Peng and Yongguo Liu, *Feature Selection and Overlapping Clustering-Based Multilabel Classification Model*, Mathematical Problems in Engineering, vol. 2018, 281489.
- [6] H. Peng, F. Long, and C. Ding, *Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005. 27(8):p. 1226–1238.
- [7] J. Nayak, et. al, *Fuzzy C-means (FCM) clustering algorithm: a decade review from 2000 to 2014 in Computational Intelligence in Data Mining—Volume 2*, L. C. Jain, H. S. Behera, J. K. Mandal, and D. P. Mohapatra, Eds., vol. 32 of Smart Innovation, Systems and Technologies, Springer, 2015. p. 133–149.
- [8] D. Mena. et al, *An overview of inference methods in probabilistic classifier chains for multilabel classification*, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2016. 36(6):p. 215–230.
- [9] F. Chamroukhi, D. Nguyen, *Model-Based Clustering and Classification of Functional Data*, 2018, arXiv: 1803.00276[cs.CR] .
- [10] Xiaohong, G., et al. *A method of vessel tracking for vessel diameter measurement on retinal images*. in Image Processing. Proceedings. International Conference on, 2001.
- [11] J. Lee and D. W. Kim, *Memetic feature selection algorithm for multi-label classification*, *Information Sciences*, 2015. 293(3):p. 80–96.
- [12] M. Lotfollahi. et al , *Deep Packet: A Novel Approach For Encrypted Traffic Classification Using Deep Learning*, 2017. arXiv:1709.02656 [cs.LG].
- [13] T. Li, et al. *Outsourced privacy-preserving classification service over encrypted data*, *Journal of Network and Computer Applications*,2018. p. 100-110.
- [14] Ehsan Hesamifard. et al, *CryptoDL: Deep Neural Networks over Encrypted Data*, arXiv:1711.05189 [cs.CR], 2017. p. 347-364,.