

A
Dissertation
On
Solving Feature Selection Problem Using Bacterial Foraging Algorithm
Submitted in Partial Fulfillment of the Requirement
For the Award of the Degree of
Master of Technology
In
Computer Science & Engineering
Submitted By
DIVYA MITTAL
University Roll No. 2K12/CSE/30
Under the Esteemed Guidance of
Dr. (Mr.) Kapil Sharma
Associate Prof., Computer Engineering Department, DTU, Delhi



2012-2015
DELHI TECHNOLOGICAL UNIVERSITY
DELHI – 110042



Department of Computer Engineering
Delhi Technological University
Delhi-110042

CERTIFICATE

This is to certify that the dissertation titled “**Solving Feature Selection Problem Using Bacterial Foraging Algorithm**” is a bonafide record of work done at **Delhi Technological University** by **Divya Mittal, Roll No. 2K12/CSE/30** for partial fulfilment of the requirements for the degree of Master of Technology in Computer Science & Engineering. This project was carried out under my supervision and has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma to the best of my knowledge and belief.

Date: _____

(Dr. (Mr.) Kapil Sharma)
Associate Professor & Project Guide
Department of Computer Engineering
Delhi Technological University

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all the people who have supported and encouraged me during the course of this project without which, this work could not have been accomplished.

First of all, I am very grateful to my project supervisor Dr. (Mr.) Kapil Sharma for providing the opportunity of carrying out this project under his guidance. I am deeply indebted to him for the support, advice and encouragement he provided without which the project could not have been a success. I am also grateful to Dr. (Mr.) OP Verma, HOD, Computer Science, DTU for her immense support. I am also thankful to my parents and spouse for being there for me at all times. Last but not the least; I am grateful to Delhi Technological University for providing the right resources and environment for this work to be carried out.

Divya Mittal

University Roll no: 2K12/CSE/30

M.Tech (Computer Science & Engineering)

Department of Computer Engineering

Delhi Technological University

Delhi – 110042

ABSTRACT

Feature selection has become important and beneficial when the amount of data and information is large. Sometimes, reduction of features can improve the quality of prediction and even be a necessary embedded step of the prediction algorithm. Further improvements in feature selection will affect a wide array of applications in fields such as biomedical, pattern recognition, machine learning, or signal processing. Bacterial Foraging Algorithm (BFA) is one of the powerful bio-inspired optimization algorithms, which attempt to imitate the single and groups of *E. Coli* bacteria. In BFA algorithm, sets of bacteria try to forage towards a nutrient rich medium to get more nutrients. In this scheme, an objective function is posed as the effort or a cost incurred by the bacteria in search of food. BFA with a Naive Bayes classifier has been presented in this work. The performance of the proposed feature selection algorithm was investigated using six benchmark datasets from different domains and was compared to three other well-known feature selection algorithms. Discussion focused on two perspectives: number of features and classification accuracy. The results showed that proposed work outperformed other algorithms in selecting lower number of features, hence removing irrelevant, redundant, or noisy features while maintaining the classification accuracy.

Index Terms - BFA algorithm, Feature selection, Naive Bayes, UCI data sets, classification

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
Figure 1:	Swimming, Tumbling and chemotactic behavior of E. coli	
Figure 2:	The clockwise and anti clockwise movements of a bacterium	
Figure 3:	Flowchart showing the complete working of the algorithm	
Figure 4:	Basic flow of proposed work	
Figure 5:	Flow chart of proposed work	
Figure 6:	Comparison on heart dataset using JRip classifier	
Figure 7:	Comparison on Vote dataset using JRip classifier	
Figure 8:	Comparison on Credit dataset using JRip classifier	
Figure 9:	Comparison on Mushroom dataset using JRip classifier	
Figure 10:	Comparison on Derm dataset using JRip classifier	
Figure 11:	Comparison on Lung dataset using JRip classifier	

LIST OF ABBERIVATIONS

ACRONYM	DEFINATION
BFO	Bacterial Foraging Optimization
BFA	Bacterial foraging Algorithm
PSO	Particle Swarm optimization Algorithm
GA	Genetic Algorithm
GPSO	Generic Particle Swarm optimization Algorithm
CPU	Central Processing Unit
SBS	Sequential Backward Selection
SFS	Sequential Forward Selection
SBFS	Sequential Forward Floating Selection
SBFS	Sequential Backward Floating Selection
MM	Max - Min
BB	Branch and bound
NB	Naive Bayes
TF	Total features
SF	Selected Features
CA	Classification Accuracy