Hydraulics & Water Resources

HYDRO 2007

Editors

P.L. Patel • B.K. Samtani

A.D. Ghare • J.N. Patel

HYDRAULICS AND WATER RESOURCES

National Conference on Hydraulics & Water Resources

HYDRO - 2007

December 21-22, 2007

Edited by

P. L. PATEL • B. K. SAMTANI
A. D. GHARE • J. N. PATEL

Organised by

Sardar Vallabhbhai National Institute of Technology, Surat
&
Indian Society for Hydraulics

In association with

L. D. College of Engineering, Ahmedabad
Maulana Azad National Institute of Technology, Bhopal
Malviya National Institute of Technology, Jaipur
Visvesvaraya National Institute of Technology, Nagpur
Indian Institute of Technology Bombay, Mumbai

Elite Publishing House Pvt Ltd

302, JMD House, 4-B, Ansari Road, Daryaganj, New Delhi - 110 002 Phone : 41004299 E-mail: rg_elite@yahoo.com

ISBN: 81-88901-32-6

Published by R. Gangadharan for Elite Publishing House Pvt. Ltd.
302, JMD House, 4-B, Ansari Road, Daryaganj, New Delhi-110002
Phone : 41004299, 26194323, 9810666503
E-mail: rg_elite@yahoo.com

Typesetting: Sai Graphic Design, 8694 D.B. Gupta Road, New Delhi-110055
Printed at Chaman Offset, New Delhi-110002

PRINTED IN INDIA

FOREWORD

The Indian Society for Hydraulics (ISH) has been set up in 1992 as a technical, education and non-profitary voluntary National Organization to encourage and foster understanding to concerned engineers and scientists engaged in various originations related to Hydraulics. One of the major objectives of the ISH is to organize National level Conference HYDRO every year for dissemination of recent contributions in the field of Hydraulics and Water Resources Engineering from scientists, academicians and Engineers. Last year HYDRO-2006 was held at Bharati Vidyapeeth College of Engineering Pune during Dec.08-09, 2006. Prior to last year conference, HYDRO conferences were held at IIT Kanpur, NIT Kurukshetra, CWPRS Pune, SIT, Tumkur, VNIT Nagpur and IIT Powai.

Sardar Vallabhbhai National Institute of Technology, SVNIT, Surat & Indian Society for Hydraulics (ISH) is organizing, National Conference on Hydraulics & Water Resources, HYDRO – 2007, during December 21-22, 2007 at SVNIT in association with L.D. College of Engineering, Ahmedabad, Maulana Azad National Institute of Technology, Bhopal, Malaviya National Institute of Technology, Jaipur, Visvesvaraya National Institute of Technology, Nagpur and Indian Institute of Technology Bombay, Mumbai.

The themes of the Conference are:

1. Water Resources & Hydrology
2. Environmental Hydraulics
3. Fluvial Hydraulics
4. Costal, Harbour and Ocean Engineering
5. Maritime Structures
6. Hydraulic Structures and Hydro Power Projects
7. Hydraulic Instrumentation
8. Soft Computing Techniques
9. Application of Geospatial Technologies
10. Fluid Mechanics

The proceedings of the Conference, HYDRO – 2007, contain technical papers, presented under the above themes and cover wide spectrum of research studies, case studies, review papers and papers related to management of water resources. It is hoped that the wealth of information covered in these papers would be very useful to the professionals working in the field of Hydraulic Engineering.

We take this opportunity to express our sincere thanks and gratitude to all the members of the Technical Committee as well as the Organizing Committee along with their volunteers, sponsorers for providing financial support and the authors to make this conference a grand success.

Dr. M. C. Deo
President-ISH
Professor & Head,
Civil Engg. Dept., IIT Bombay
Chairman, HYDRO - 2007

Dr. B. K. Samtani
Organising Secretary
HYDRO-2007
SVNIT, Surat

Dr. P. D. Porey
Director
SVNIT, Surat

Sponsorers

- Technical Education Quality Improvement Programme, New Delhi
- All India Council for Technical Education, New Delhi
- Department of Science and Technology, New Delhi
- Council for Scientific and Industrial Research, New Delhi
- University Grants Commission, New Delhi
- National Institute of Technology, Agartala
- M N National Institute of Technology, Allahabad
- National Institute of Technology, Calicut
- National Institute of Technology, Durgapur
- National Institute of Technology, Hamirpur
- National Institute of Technology, Jamshedpur
- National Institute of Technology, Jallandar
- National Institute of Technology, Kurukshetra
- National Institute of Technology, Patna
- National Institute of Technology, Raipur
- National Institute of Technology, Rourkela
- National Institute of Technology, Silchar
- National Institute of Technology, Srinagar
- National Institute of Technology, Surathkal
- National Institute of Technology, Trichi
- National Institute of Technology, Warangal
- Indian Institute of Technology, Roorkee
- Indian Institute of Technology, Kanpur
- Indian Institute of Technology Madras, Chennai
- G H Raisoni College of Engineering, Nagpur
- National Thermal Power Coporation Ltd., New Delhi
- Reliance Energy Limited, Noida
- ESSAR Power Gujarat Ltd., Mumbai
- Ambuja Cements Ltd., Ahmedabad
- DHI (I) Water and Environment (P) Ltd., New Delhi
- Water and Power Consultancy Services (I) Ltd., Gurgaon
- Rachna Pipes, Surat
- Environmental Measurements and Control, Kochi

Organising Committee

Director & Patron	:	**Dr. P.D. Porey**, Director, SVNIT, Surat.
Chairman	:	**Prof. M.C. Deo**, President ISH, IIT, Bombay
Co-chairman	:	**Shri P.B. Deolalikar**, Vice President ISH **Dr. N.J. Mistry**, Prof & Head, CED, SVNIT
Organising Secretaries	:	**Prof. B.K. Samtani**, SVNIT **Dr. S. Balakrishna**, ISH, CWPRS, Pune
Joint Secretaries	:	**Prof. J.N. Patel**, SVNIT **Prof. M.B. Dholakia**, LDCOE, Ahmedabad **Prof. Sushil Kumar**, MANIT, Bhopal **Prof. Rohit Goyal**, MNIT, Jaipur **Shri A.D. Vasudeo**, VNIT, Nagpur
Registrar	:	**Shri H.A. Parmar**, Registrar, SVNIT, Surat.

MEMBERS

Prof. K.C. Pandey
MANIT, Bhopal

Prof. Sunil Sarangi
NIT, Rourkela

Prof. Y.V. Rao
NIT, Warangal

Prof. K.M. Shah
LDCOE, Ahmedabad

Prof. Baldev Sethia
NIT, Kurukshetra

Prof. G.L. Asawa
IIT, Roorkee

Prof. Rajneesh Srivastava
AICTE, New Delhi

Mrs. V.V. Bhosekar
CWPRS, ISH, Pune

Prof. N.C. Shah
SVNIT, Surat

Prof. R.P. Dahiya
MNIT, Jaipur

Prof. M.N. Bandopadhyay
NIT, Kurukshetra

Prof. Wanchoo
NIT, Srinagar

Prof. H.K. Raval
SVNIT, Surat

Dr. C.T. Mahajan
NPIU, New Delhi

Mrs. V.M. Bendre
CWPRS, Pune

Prof. Swadesh Gupta
AICTE, New Delhi

Shri P.B. Mehendale
CWPRS, ISH, Pune

Dr. Rajesh Srivastava
IIT, Kanpur

Dr. A.D. Ghare
SVNIT, Surat

Prof. G.R.C. Reddy
NIT, Calicut

Prof. S.S. Gokhale
VNIT, Nagpur

Prof. M. Chidambaram
NIT, Trichi

Prof. S. Mohan
IIT, Madras

Prof. T.I. Eldho
IIT, Bombay, Mumbai

Prof. H.K. Mishra
COE, Jabalpur

Shri Milind Kulkarni
DST, New Delhi

Prof. L.M. Gupta
VNIT, Nagpur

Prof. P.L. Patel
SVNIT, Surat

Sub Committees

Reception Committee
Dr. N.C. Shah, Prof., CED (Convener)
Shri H.A. Parmar, Registrar
Dr. N.J. Mistry, HOD, CED
Dr. B.K. Samtani, Prof., CED
Dr. G.N. Mehta, Prof., ASHD
Dr. J.A. Desai, HOD, AMD
Dr. D.P. Vakharia, Dean (SC)
Dr. Smita Jauhari, Lecturer, ASHD

Registration Committee
Dr. J.N. Patel, Prof. CED (Convener)
Dr. S.A. Channiwala, Prof., MED
Dr. H.K. Raval, Prof., MED
Dr. D.R. Patel, Prof., COED
Dr. ZVP Murthy, Prof., CHED
Dr. M.A. Zaveri, AP, COED
Dr. P.R. Tailor, Prof., MED
Shri S.N. Desai, AP, AMD
Shri A.K. Desai, Lecturer, AMD
Shri P.G. Agnihotri, Lecturer, CED
Ms. Manisha Chauhan, Lecturer, EED
Ms. N.D. Jariwala, Lecturer, CED
Shri Y.D. Patil, Lecturer, AMD
Shri G. D. Kale, Lecturer, CED

Finance Committee
Shri H.A. Parmar, Registrar (Convener)
Dr. B.K. Samtani, Prof., CED
Dr. H.S. Patil, Prof., AMD
Dr. P.A. Parikh, Prof., CHED
Dr. Jigisha Parikh, AP, CHED
Shri S.C. Rathod, Dy. Registrar (A/c.)
Shri R.C. Mody, Sr. Clerk, CED

Press, Publicity & Invitation Committee
Dr. P.D. Porey, Director (Convener)
Dr. B.K. Samtani, Dean (Exam)
Dr. N.C. Shah, Dean (UG)
Shri H.A. Parmar, Registrar
Dr. R.V. Rao, HOD, MED
Dr. P.L. Patel, Dean (PG)
Prof. A.J. Shah, Lecturer, AMD
Shri Kunwar Yadav, Lecturer, CED
Shri K.P. Solanki, I/C. PA to Director

Editorial Committee
Dr. P.L. Patel, Dean (PG) (Convener)
Dr. B.K. Samtani, Dean (Exam)
Dr. J.N. Patel, Professor, CED
Dr. A.D. Ghare, AP, CED
Dr. Rohit Goyal, MNIT, Jaipur
Prof. M.B. Dholakia, LDCOE, A'bad
Shri G.D. Kale, Lecturer, CED
Shri G.R. Veshmawala, Lecturer, AMD
Shri T.B. Ghosh, Librarian

Accommodation Committee
Dr. D.P. Vakharia, Dean (SC) (Convener)
Dr. D.R. Patel, Prof., COED
Dr. N.K. Datta, AP, Phy. Edn.
Ms. M. Chakraborty, AP, CHED
Shri S.M. Yadav, Lecturer, CED
Dr. A.K. Rai, Lecturer, ASHD
Dr. Amit Solanki, Lecturer, AMD
Dr. L.K. Saini, Lecturer, ASHD
Shri A. Mandaloi, Lecturer, ECED
Shri K.S. Bhavsar, Jr. Clerk, A/Cs.
Shri P.G. Naidu, Jr. Clerk, B&C

Catering Committee
Prof. A.R. Jariwala, HOD, EED (Convener)
Prof. H.D. Desai, AP, MED
Prof. B.K. Samtani, Prof., CED
Shri H.A. Parmar, Registrar
Dr. H.J. Nagarsheth, AP, MED
Dr. S.G. Shah, Medical Officer
Dr. G.J. Joshi, Chief Hostel Warden
Shri U.P. Rao, Lecturer, COED

Technical Session Committee
Dr. B.K. Samtani, Dean (Exam) (Convener)
Dr. P.L. Patel, Prof., CED
Dr. J.A. Desai, HOD, AMD
Dr. H.K. Raval, Prof., MED
Prof. R.T. Patel, Lecturer, AMD
Dr. A.D. Ghare, AP, MED

Venue Committee

Dr. N.J. Mistry, HOD, CED (Convener)
Dr. D.R. Patel, Dean (A&RG)
Dr. H.B. Naik, Prof., MED
Prof. D.B. Raijiwala, AP, AMD
Dr. D.V. Bhatt, AP, MED
Shri J. Banerjee, AP, MED
Shri M.N. Bhusavalwala,
Mrs. A.K. Khambhete
Shri J.N. Desai, Sr. Clerk, A/Cs.
Shri P.B. Patel, Hamal, Library

Transport Committee

Dr. C.D. Modhera, AP & Supdt. Engr. (Convener)
Shri S.N. Gaderia, AP, MED
Prof. V.M. Daiya, AP, MED
Shri S.M. Yadav, Lecturer, CED
Shri B.M. Sutaria, Lecturer, MED
Shri M.B. Maisuria, Lecturer, MED
Shri R.P. Gohil, Lecturer, COED
Shri S.R. Arya, Lecturer, EED
Shri A.P. Patil, Lecturer, MED
Shri Rakesh Maurya, Lecturer, EED
Shri Ranjit Roy, Lecturer, EED
Shri Piyush N. Patel, Lecturer, ECED
Shri S.R. Goyal, Lecturer, COED
Shri Dinesh Singh, Lecturer, MED
Shri B.N. Tandel, Lecturer, CED
Shri Ashish Dhamania, Lecturer, CED
Shri A.M. Pardhi, Lecturer, AMD
Shri M.L. Patel, Jr. Clerk, E&S

Volunteer Committee

Dr. A.D. Ghare, AP, CED (Convener)
Dr. M.N. Mehta, HOD/ASHD
Dr. A.K. Mungrey, Lecturer, CHED
Shri D.B. Gohil, Lecturer, MED
Ms. K.C. Maheria, Lecturer, ASHD
Ms. Dimple Shah, Lecturer, ASHD
Shri S.R. Patel, Lecturer, CHED

Administrative Support Committee

Shri H.A. Parmar, Registrar (Convener)
Shri M.N. Patel, Dy. Registrar (Acad)
Shri A. Mahto, Lecturer, MED
Shri Achchhelal, Lecturer, MED
Shri M.V. Tadvi, Lecturer, MED
Shri G.A. Mistry, ‘E’ Section
Mrs. A.T. Elavia, Head Clerk, Hostel
Shri P.P. Vankar, Draftsman, MED
Shri R.D. Ruparel, Jr. Clerk, A/cs.

Infrastructure Committee

Dr. R.A. Christian , Dean (P&D) (Convener)
Prof. JEM Macwan, CED
Shri N.N. Patel, CED
Prof. B.R. Taunk, HOD/ECED
Dr. K.P. Desai, AP, MED
Shri S.M. Yadav, Lecturer, CED
Shri K.A. Chauhan, Lecturer, CED
Shri Pranav Darji, Lecturer, EED
Shri D.A. Patel, CED
Shri S.C. Shoemaker, AAEE, E&S
Shri J.N. Desai, Sr. Clerk, A/Cs.
Shri J.C. Shah, Sanitary Inspector, E&S
Shri B.C. Talavia, Workshop

Ladies Committee

Mrs. Prof. N.Y. Desai, AP, ECED (Convener)
Mrs. Prof. V.A. Shah, AP, EED
Mrs. R. Chudamani, AP, EED
Dr. Mrs. M. Mukhopadhyay, AP, CHED
Mrs. N.D. Jariwala, Lecturer, CED
Mrs. J.V. Menghani, Lecturer, MED
Ms. Chandani Gor, Lecturer, EED
Ms. J. Solanki, Lecturer, CHED
Ms. Premalata Kumari, Lecturer, ASHD

Contents

ENVIRONMENTAL HYDRAULICS

Classification of Rainy Days Using *SVM*

Anirudh Vemula* and Umesh C. Kothyari**

ABSTRACT

The process of occurrence of rainfall involves significant variations both temporally as well a spatially. In this study, an attempt is made to model occurrence of rainy days using support vector machines (*SVM*). The *SVM* is a novel learning machine that is based on statistical learning theory. The *SVM* involves determination of optimal separating hyper-plane in the input space. Kernel functions are employed to map the non-linear input space to a linear output space of higher dimension. Linear and Gaussian kernels were used for constructing the *SVM*. This study indicated that the classification accuracy of the *SVM* model is greatly influenced by model parameters namely cost parameter (C) and gamma parameter (γ).

INTRODUCTION

Most of the hydrological applications use rainfall amount rather than the rainy days as the inputs. Some physical systems are however affected more by the occurrence of rainfall or rainy days rather than the rain amount. The occurrence of rainfall rather than its amount has been used as a significant input variable in water demand modeling and management (Jain *et al*., 2001). Therefore the modeling of the rainfall occurrence is also important.

Considerable research has been conducted during the past in developing the mathematical models for simulation of the process of rainfall. However, the process of formation, mechanism, occurrence and its temporal as well as spatial distribution involve a rather complex Physics that is not completely understood as yet. The review of studies employing various techniques for rainfall modeling can be found in Belloc (1980), Georgakakos and Hudlow (1984). French *et al.* (1992) were amongst the first to employ the machine learning technique of artificial neural network *ANN* for rainfall forecasting. *ANN*s were developed for rainfall forecasting by transformation of a rainfall intensity field from the current state to a 1 h ahead state. The *ANN* was shown by French et al. to perform forecasting at a high level of accuracy.

Similar studies however, have not been conducted as yet by using the support vector machines *SVM* models. The *SVM* is one among the novel learning machine techniques that is based on

* Undergraduate student, Department of Civil Engineering, Indian Institute of Technology, Roorkee Email : ani19uce@iitr.ernet.in

** Professor Department of Civil Engineering,Indian Institute of Technology, Roorkee, E-mail: umeshfce@iitr.ernet.in

statistical learning theory similarly as the *ANN*. Although *SVM* have been used in a range of problems in the other fields of studies, the present investigation study is perhaps the first attempt to apply them in hydrological investigation on the classification of the rainy days.

Despite of its very good success in other fields of study, there are only a few applications of *SVM* in the field of hydrology (Asefa et al., 2006). Dibike et al. (2001) applied *SVM* in remotely sensed image classification and regression for rainfall-runoff modeling investigations. Comparison between *SVM*, *ANN* and conceptual rainfall-runoff models at three catchments were made by them which showed the *SVM* to perform the best way amongst the methods used. Liong and Sivapragasam (2002) used flood stages at a number of locations in a river on its up stream to predict the river stage at the downstream location. Liong and Sivapragasam also reported a superior *SVM* performance compared to *ANN* performance in forecasting flood stage. Asefa and Kemblowski (2002) used *SVM* to simulate the groundwater flow and contaminant transport. Yu et al. (2005, 2006) have applied the *SVM* more recently in flood stage and discharge predictions. Asefa et al. (2006) have shown applications of *SVM*s in predicting both seasonal and hourly stream flows.

In the present study, the attempt is made to classify rainy days by using the *SVM*. The emphasis is given to the occurrence of rain, rather than the magnitude of rainfall itself. The occurrence of rainy day is assumed herein to be a function of the other climatic variables on that particular day. The proposed models were developed and evaluated using the meteorological data observed at National Institute of Hydrology, Roorkee, India. The performance of all the proposed models has been evaluated in terms of their classification accuracy.

PROPOSED METHODOLOGY

The *SVM* were first suggested by Vapnik (1995) and have recently been used in a range of problems such as pattern recognition, bioinformatics and text categorization. They are based on structural risk minimization (*SRM*) principle; theoretically minimizing the expected error of a learning machine and reducing the problem of overfitting. However the *SVM* is used herein by considering a binary classification problem:

Let (x_i, y_i) be a training set of instance-label pairs, $i = 1,...,m$ where $x_i \in R^n$ and $y_i \in \{-1,1\}$. Here R represents the set of real numbers and n is the number of dimensions. Hence for a linearly separable case, the data will be classified by the following separating hyper-plane

$$y_i(\langle w.x_i \rangle + b) - 1 \geq 0 \;\; \forall \; i = 1,...,m \tag{1}$$

where w is the normal vector and b is the intercept on Y axis. To maximize the margin width of the separating hyper-plane, we need to solve the following optimization problem:

$$\underset{w,b}{Min} \frac{1}{2} w^T w \quad \textit{subject to}: \; y_i(\langle w.x_i \rangle + b) - 1 \geq 0 \tag{2}$$

The solution of Eq. (2) is also called the maximum margin or hard margin classifier (see Fig. 1), which however, does not have a solution for non-separable data. Further, even for the separable case, we can increase the margin width. Thus, for better classification or for non-separable case, the above equation (Eq. 2) can be solved by introducing non negative slack variables $\xi_i \geq 0$, $i = 1,...,m$ such that $\langle w.x_i \rangle + b \geq 1 - \xi_i$ for $y_i = +1$ and $\langle w.x_i \rangle + b_i \geq 1 - \xi_i$ for $y_i = -1$.

All the computations are carried herein on via kernel function in input space. There are several possibilities for the choice of kernel function, including linear, polynomial, sigmoid, splines, Gaussian kernel and in fact *any* function that satisfies *Mercer's theorem* (Vapnik, 1998, p.136)

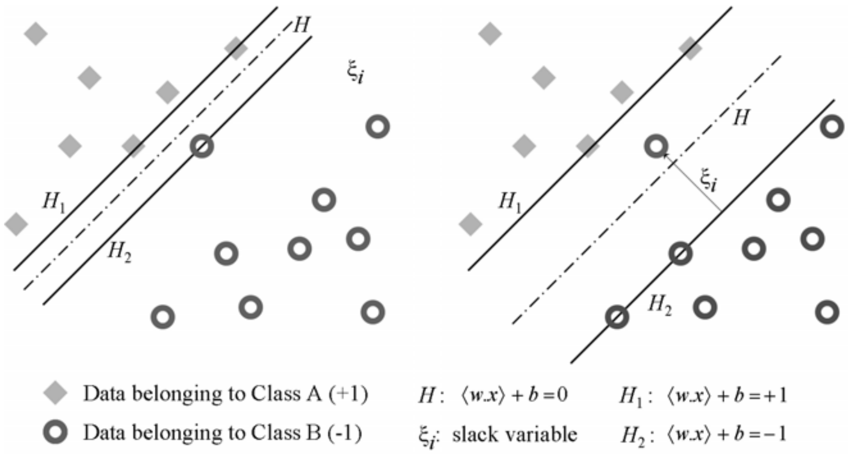


Fig. 1: Hard margin classifier **Fig. 2: Soft margin classifier**

can be used as a kernel. However, the Gaussian kernel has been widely used since it possesses many well-known good properties. Moreover, the linear kernel is a special case of Gaussian kernel (Keerthi and Lin, 2003). In this study Gaussian kernel is used to map the input data into higher dimensional feature space, which is given by:

$$K(x, y) = \exp(-||x-y||^2/(2\sigma^2)) \tag{10}$$

$$K(x, y) = \exp(-^{3}||x-y||^2) \tag{11}$$

where, σ is the called the spread parameter of Gaussian kernel, which can be adjusted to control the expressivity. The Gaussian kernel has a localized and finite response across the entire range of predictors and it nonlinearly maps the training data into a possibly infinite dimensional space. Thus, it can effectively handle the situations when the relationship between predictors and the target and is non-linear. Moreover, it is computationally simple than polynomial kernel, which has more parameters.

Genetic Algorithms

Genetic Algorithms (*GA*) are search algorithms, based on the mechanics of natural selection and natural genetics (Goldberg, 1989). They are an adaptive optimization methodology, and are a promising alternative to conventional heuristic methods. Although stochastic in nature, they are directed and also exploit the prior information to guess new search points. *GA* generates successive populations of alternate solutions that are represented by a chromosome, *i.e*. a solution to the problem, until acceptable results are obtained. The search is guided by the objective or fitness function which determines the quality or fitness of a solution in the evaluation step. The three basic genetic operators are – reproduction or selection operator, cross over and mutation. The search concludes when the termination conditions are satisfied.

Two models have been proposed in this study: *GA optimized linear SVM classifier* and *GA optimized Gaussian kernel classifier*. For the linear classifier, the chromosome consists of the only the cost parameter *C*. For the Gaussian kernel classifier, the chromosome consists of two parts: *C* and *γ*. The fitness function of these models is defined such that higher accuracy implies higher fitness value; the accuracy of classification being defined as the ratio of number of correct classifications to the total classifications performed.

DATA, RESULTS AND DISCUSSIONS

In the present study, the meteorological data of Roorkee observed at weather station of the National Institute of Hydrology, Roorkee, India were used to evaluate the classification accuracy of the proposed schemes. The data consisted of daily values of maximum and minimum temperature, pan evaporation, relative humidity, wind speed and corresponding daily rainfall for 2061 days spanning over the time period between January 1, 2000 and December 31, 2005. The range of pertinent climatic variables of the data is listed in Table 1.

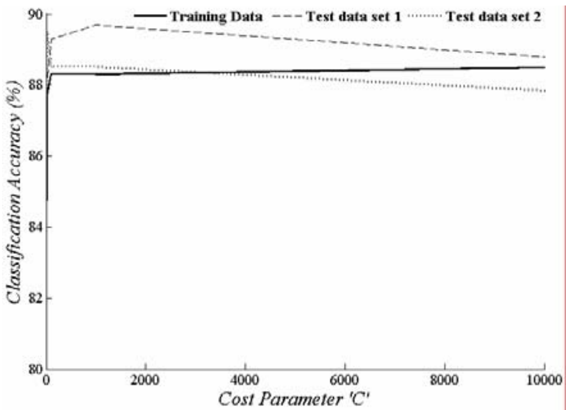
Table 1 Range of data used

Data set.	Daily Rainfall (mm)	Relative humidity (per cent)	Wind speed (Km/hr)	Maximum Temperature (ºC)	Minimum Temperature (ºC)	Pan Evaporation (mm)
Training	0 to 162.2	29 to 100	0 to 8.3	9.0 to 44.2	2.3 to 31	0.0 to 11.3
Test set 1	0 to 116.4	8 to 100	0 to 6.0	9.0 to 40.0	2.5 to 36.5	0.0 to 7.2
Test set 2	0 to 119.2	35 to 100	0 to 6.0	10.5 to 43.7	1.6 to 28.8	0.0 to 10.8

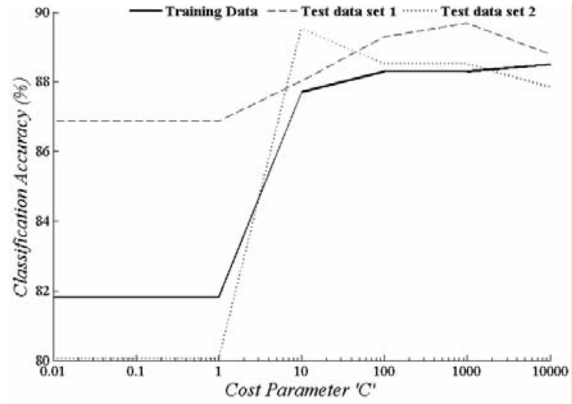
The data was scaled linearly as described by Hsu *et al*., (2004) such that the minimum is represented by zero and maximum by unity. This prevents the attributes in greater numeric ranges to dominate those in smaller numeric ranges. Moreover, it also avoids numerical difficulties during the computations especially for the inner products (for the kernel values). The data set was then randomly split into 3 sets: a training set of 1000 data points and two test data sets of equal length to ensure the validity of the results and to check overfitting of the proposed models, the *ANN* based models in particular. Finally, a threshold limit of 1 mm of rainfall was used to define a rainy day and also to prevent noise in the input data. Daily values of; maximum and minimum temperature, pan evaporation, relative humidity and wind speed formed the model input data in present application. The corresponding days rainfall was used to decide the target variable *i.e.* rainy day. As mentioned already, the day for which rain occurred exceeded 1mm was termed as the 'rainy day'. The proposed models were constructed accordingly and their performance was studied in classifying the rainy days. The results are reported in terms of the classification accuracy that was defined as the ratio of number of days having correct classification of the rainy days to the total number of days in the particular data set, expressed in terms of percentage.

The proposed *SVM* models were implemented on MATLAB development environment 7.0 by applying the LIBSVM, originally developed by Chang and Lin (2001). The classification accuracy for the linear and Gaussian kernels was plotted against C and γ to study the variation. The optimization of the *SVM* parameters was performed using the MATLAB 'Genetic Algorithm and Direct Search Toolbox 1.0.2'. Finally, the values of the parameters were chosen such that the training accuracy was maximized and to overfitting is prevented as the results validated using the two test data sets.

It may be observed from the Fig. 3 (a) that the performance of the linear *SVM* classifier increases rapidly for the training data while C is small and subsequently it increases slowly with the increase in the cost parameter. However, the performance of the test data sets increases rapidly while C is small, until it reaches a maximum and then it falls. From the Fig. 3(b) it can be seen that the optimal value of C lies somewhere between 1 and 10. The value of optimum C as determined



(a) Linear Scale



(b) Log scale

Fig. 3: Variation of classification accuracy of linear *SVM* classifier with cost parameter

Table 2: Results for *GA* optimized linear classifier

C	Classification accuracy (%)		
	Training	Test data set 1	Test data set 2
2.2383	82.2	88.8	85.8
2.4221	82.9	89.6	85.1
2.5145	83.2	89.3	85.5
3.3131	85.6	89.9	87.2
3.9269	86.8	89.2	89.2
5.6127	87.1	88.2	89.5

using *GA* on training data alone is 5.6127, and the corresponding classification accuracies being 87.1%, 88.15% and 89.5% for the training and test data sets respectively (Table -2).

It was noticed in Fig. 4 that as the value of the penalty parameter C and γ increase, the performance of the Gaussian kernel *SVM* classifier for training data also increases. However, the performance for testing data sets decreases with the increase in the value of these parameters.

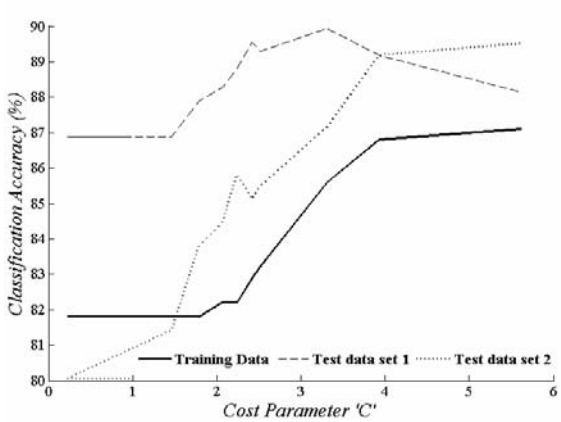


Fig. 4: Performance of *GA* based linear *SVM* classifier

This may be attributed to the overfitting of the classifier to the training data. Moreover, it was also seen that there are '*valleys*' where the performance of the Gaussian kernel *SVM* classifier fell drastically for the testing data sets. Thus, the choice of the parameters must be so fixed such that the performance of the classifier is maximized for any data set.

Table 3: Experimental results for *GA* optimized Gaussian kernel *SVM* classifier

Cost parameter	Gamma parameter	Classification accuracy (%)		
		Training Data	Test data set 1	Test data set 2
0.5915	0.7621	85.6	89.6	85.5
0.9342	0.6154	86.2	91.0	86.8
0.8729	0.7382	86.4	90.3	87.5
0.9669	0.9355	86.8	90.6	87.8
2.2496	1.343	87.8	89.8	88.2
1.4831	2.2482	87.5	89.5	88.5
3.2633	2.9208	88.0	89.3	88.9

Therefore, the optimum values of these parameters as found using *GA* are $C = 3.2633$ and $\gamma = 2.9208$ and the classification accuracies for the training and testing data sets are 88%, 89.3% and 88.85% respectively (Table 3). It can also be noted by comparison between Table 2 and Table 3 that the performance of the Gaussian kernel model exceeds the performance of the linear classifier, clearly showing its superior classification ability.

CONCLUSIONS

The occurrence of rainy day can be more important factor than the magnitude of rainfall itself for certain real life applications. Many investigators have developed models for predicting rainfall magnitude but there is also a need to develop models only for occurrence of rainy days. This paper presents the results of an investigation on development of model for predicting rainfall occurrence using *SVM*. To construct and evaluate the proposed schemes, the daily meteorological data of Roorkee, India were used, which consisted of daily values of maximum and minimum temperature, evaporation, humidity, wind speed and corresponding daily rainfall. The results produced by the SVM modelling have been considered as satisfactory due to the given complexities present in the process of occurrence of the rainy days.

REFERENCES

1. Asefa T. and Kemblowski M.W. (2002). Support vector machines approximation of flow and transport models in initial groundwater contamination network design, *Eos. Trans. AGU.,*vol. **83**.
2. Asefa, T., Kembloski, M., McKee, M. and Khalil, A. (2006). Muti-time scale stream flow predictions: The support vector machine approach. *Journal of Hydrology*, vol.318(1-4), 7-16.
3. Belloc, A. (1980). Operational methods of quantitative precipitation forecasts for hydrological purposes and possibilities of an inter-comparison, WMO, Geneva, pp6.
4. Chang, C.C., and Lin, C. (2001). LIBSVM : a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

5. Hsu, C., Chang, C. and Lin, C. (2004). A practical guide to support vector classification. Technical Report, National Taiwan University. Available at http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.
6. Dibike, B.Y., Velickov, S., Solomatine, D. and Abbott, M.B.(2001). Model induction with support vector machines: introduction and applications, *J. Comput. Civil Eng.*,vol.**15**(3), 208–216
7. French, M.N., Krajewski, W.F. and Cuykendall, R.(1992). Rainfall forecasting in space and time using a neural network, *Journal of Hydrol.*,vol.137,1-31.
8. Georgakakos, K.P. and Hudlow, M.D. (1984). Quantitative precipitation forecast techniques for use in hydrological forecasting, *Bull Amer. Meteor. Soc.,* vol.65(11), 1186-2000.
9. Goldberg, D.E.(1989). Genetic Algorithms in Search, Optimization, and Machine Learning. *Reading*, MA: Addison-Wesley.
10. Haykin, S., 2003. Neural Networks: A comprehensive foundation. *Fourth Indian Reprint, Pearson Education*, Singapore, pp. 842.
11. Hsu, C.W. and Lin, C.J. (2002). A simple decomposition method for support vector machine. *Machine Learning,* vol.46(1–3), 219–314.
12. Jain,A., Varshney, A.K., and Joshi, U.C. (2001). Short-term water demand forecast modeling at IIT Kanpur using artificial neural networks, *Water Resour. Mgmt.*,vol.15(5),299-321.
13. Keerthi, S.S., Lin,C.J. (2003). Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation,* 15 (7), 1667–1689.
14. Liong, S.Y. and Sivapragasam, C. (2002). Flood stage forecasting with SVM, *AWRA* **38** (2002) (1), pp. 118–173.
15. LaValle, S.M. and Branicky, M.S. (2002). On the relationship between classical grid search and probabilistic roadmaps. *International Journal of Robotics Research*, vol. 23(7–8), 673–692.
16. Trevor, H., Saharon, T., Robert, T. and Ji, Z. (2004). The entire regularization path for the support vector machine, *Journal of Machine Learning Research,* 51391-1415.
17. Valdes, J.B. and Sastri, T. (1989). Rainfall intervention analysis for on-line applications, *Jour. Water Resour. Plng. & Mgmt.*, ASCE, vol.115(4), 397-415.
18. Vapnik, V.N. (1998). Statistical Learning Theory. Wiley, New York.
19. McCulloch, W.S. and Pitts, W. (1943). A logical calculus of the ideas immanent in neurons activity, *Bull. Math. Biophys*., vol. 5, pp. 115–133.
20. Yu, P.S., Chen, S.T. and Chang, I. F. (2005). Flood stage forecasting using support vector machines, *Geophysical Research*, vol. 7(04176).
21. Yu, P.S., Chen, S.T. and Chang, I.F. (2006). Support vector regression for real time flood forecasting. *Journal of Hydrology*, vol. 328(3-4), 704-716.