

A COMPREHENSIVE STUDY ON DIFFERENT PATTERN RECOGNITION TECHNIQUES

By

NAVNEET KUMAR KASHYAP *

B.K. PANDEY **

H.L. MANDORIA ***

ASHOK KUMAR ****

* PG Scholar, Department of Information Technology, G.B.P.U.A & T, Pantnagar, India.
, ** Assistant Professor, Department of Information Technology, G.B.P.U.A & T, Pantnagar, India.
*** Professor and Head, Department of Information Technology, G.B.P.U.A & T, Pantnagar, India.

ABSTRACT

Pattern Acceptance has been admired because of its advancement in the appliance areas. The applying breadth includes medicine, communications, automation, aggressive intelligence, abstracts mining, bioinformatics, certificate classification, accent recognition, business and abounding others. In this analysis, cardboard assorted approaches of Arrangement Acceptance has been presented with their pros-cons, and the appliance specific archetype has been confirmed. From the base of the survey, arrangement acceptance techniques could be categorized into six parts. Such owning techniques include Neural Network scheme, Statistics Techniques, Template Matching, Hybrid versions and Fuzzy Model.

Keywords: Statistical & Structural Pattern Recognition, Pattern Recognition Techniques, Pattern Recognition, Fuzzy Model, Neural Network Scheme, Hybrid Versions.

INTRODUCTION

Perceiving the items and the encompassing the environment is an insignificant errand for the people. In any situation, if the objective of actualizing it comes falsely, then it turns into an extremely complex assignment. Design Recognition gives the answer for different issues from discourse acknowledgment, face acknowledgment to the characterization of written by hand characters and restorative analysis.

The different application regions of example acknowledgment resemble bioinformatics, report order, picture examination, information mining, modern robotization, biometric acknowledgment, remote detecting, written by hand content investigation, restorative determination, discourse acknowledgment, GIS and some more. Closeness between every such type of applications is that for an answer observing methodology, highlights must be extricated and broken down afterward for acknowledgment and grouping reason.

Three procedures come about in the circumstance

recognition assignment. The initial step is obtaining the information. Information obtaining is the procedure of changing over the information from one structure (discourse, character, pictures and others.) into another, which ought to be adequate to the registering system for further handling. Information obtaining is for the most part performed by sensing element, digitizing device and scanners. The second step is information examination. After information obtaining, the assignment of examination starts. Amid the information investigation step, the finding out about the information happens and data is gathered about the distinctive occasions and routine classes accessible in the information [7].

This data or learning about the information is implemented for further preparing. The third step utilized for pattern exposure is characterization. Its motivation is to choose the classification of new information on the premise of learning had gotten from the information examination process [24],[25]. The information set exhibited to a Pattern Recognition framework is isolated into two sets: the preparing set and the testing set. Framework gains from the preparing set and productivity of the framework are

checked by introducing the testing set to it [13]. The execution of the example acknowledgment systems is impacted by principally three components:

- Measure of information
- Innovation used (method)
- Developer and the client.

1. Objective of the Study

The main Objective of this comprehensive study is discussing the different approaches of pattern recognition. The evaluation work in pattern acknowledgment is to create frameworks with the capacity of taking care of gigantic measures of information. The different models settled on example acknowledgment are:

Factual methods, Structural methods, format Matching, Neural Network dependent techniques, Fuzzy models and Hybrid systems. Block diagram of a pattern recognition system is shown in Figure 1.

2. Pattern Recognition Models

The models decided on pattern detection can be sorted into various classifications relying on the technique utilized for information evaluation and organizing. Models can be freely or conditionally used to perform pattern recognition [17]. The diverse models utilized for pattern identification errand are discussed below.

2.1 Statistical Model

In the Statistical technique for Pattern identification, every routine is represented as far as components. Components are picked in a manner such that a distinctive pattern possesses a non-covering highlight space. It perceives the probabilistic nature, both of the data we choose to handle and of the structure in which we ought to present it. It functions admirably when they chose the highlights that lead to the highlight spaces,

which clusters in a distinguished way, i.e. there is a legitimate interclass separation.

In the wake of breaking down, the likelihood circulation of a routine or pattern having a place with a class has a choice limit which is resolved [3], [4]. Here the pattern is anticipated to the pre-preparing operations to make them reasonable for preparing purposes. Elements are chosen after breaking down the training designs. The framework learns and adjusts the unknown hidden pattern as appeared in Figure 2. Test pattern is connected to check the reasonableness of the framework to perceive patterns. Highlight estimation is done while testing, then these element qualities are displayed to the educated framework and along these lines, an arrangement is performed [19]. At the point when contingent likelihood thickness circulation is known, parametric arrangement plans are utilized, generally a nonparametric categorization plan should be used. Different choice guidelines are there to decide the choice limits like Bayes Decision Rule, Optimal Bayes Decision Rule, The Maximum Likelihood Rule, Neyman-Pearson lead, and MAP standard [2],[1]. As highlight spaces are apportioned, the framework gets to be noise insensitive in this way, if there should arise an occurrence of noisy patterns. The decision of measurable model is an excellent arrangement. It all depends on whether the strategy selected is managed or unsupervised statistical approach procedure can be classified as Discriminant Analysis and Principal Component Analysis [1].

Discriminant Analysis is a regulated method in which we accomplish for dimensionality elimination. Here, straight mix of elements is used to perform the order operation. For every pattern class, a Discriminant capacity is characterized which performs the categorization functionality [8].

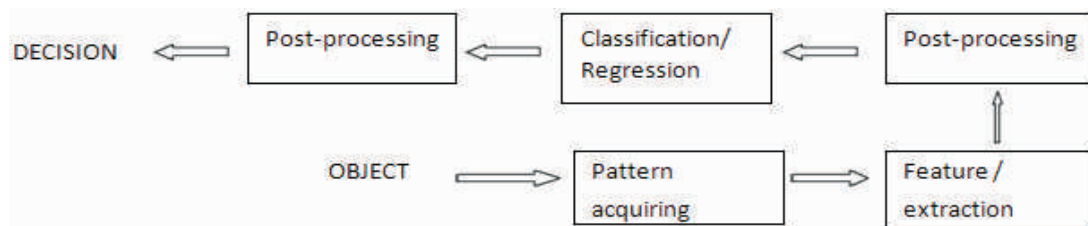


Figure 1. A Pattern Recognition System



Figure 2. Statistical Pattern Recognition Model

There is not an all around characterized guideline with respect to the type of Discriminant capacity like MDC (Minimum Distance Classifier) that utilizes one reference point for each and every class, and the Discriminant capacity results in a distance classifier from the obscure vectors to these focuses, and then again closest neighbor classifier utilizes the set of focus for every class.

There are various of Discriminant Analysis techniques that are utilized based upon the application and framework prerequisite, for example, Null-LDA (N-LDA), LDA (Linear Discriminant Analysis), 2D-LDA (Two Dimensional Linear Discriminant Analysis), FDA (Fisher Discriminant Analysis), 2D-FDA (Two Dimensional Fisher Discriminant Analysis) [4],[5]. In LDA, the highlight set is gotten by direct blend of unique components. Intra-class separation is minimized and additionally augmentation is done between the class separation, to acquire the ideal results. LDA experiences little specimen size (SSS) issue.

In FDA, the proportion of change in between the classes to variations in the intra-classes characterize the detachment between classes. Here, class diffuse is amplified and intra-class disseminate is minimized to get the ideal results [8]. FDA methodology is a mix of PCA and LDA. 2D-LDA maintains a strategic distance from the little example size (SSS) issue connected with 1D-LDA. Here, the lattices of information are figured to frame the element vector.

Hint of interclass diffuse matrix is optimized, while hint of intra-class disseminate matrix is decreased to get the ideal results in 2D-LDA. When compared with 1-D LDA; 2D-FDA gives non-particular interclass and intra-class lattices. Ching et al. [9] proposed that the invalid space spread over by the eigenvectors of the intra-class dissipate lattices having zero Eigen values that contains profoundly segregating data.

An LDA strategy in the invalid space of intra-class dissipate network is N-LDA, which includes taking care of the Eigen value issue for a nationwide matrix. Principal Component Analysis (PCA) or Karhunen-Loeve development is a multi-component unsupervised strategy in which, the authors approach for dimensionality lessening [10]. Utilizing PCA, patterns are distinguished in the information and these patterns decide the comparability measure [11]. In PCA, Eigen vectors with the biggest Eigen qualities are registered to shape the element space.

PCA is firmly identified with Factor evaluation [11]. Kernel PCA (Principal Component Analysis) is an answer for nonlinear element extraction [12]. Different not-linear element extraction procedures are Kohonen highlight Map & MDS (Multidimensional Scaling) [14]. The application regions of PCA incorporate the graphically untrustworthy temperamental patterns. Discriminant evaluation is more proficient when compared with PCA, as far as precision and the time slipped by [15].

2.2 Structural Model

When we ran over patterns with solid inviolable structures, analytical techniques give vague outcomes, since highlight extraction wrecks the indispensable data concerning the important framework of a pattern. Hence, in the complex pattern recognition issues, similar to an acknowledgment of multidimensional items, it is necessary to receive a various leveled framework (hierarchical), where an example is thought to be comprised of basic sub-designs, which are further made out of less difficult sub patterns [16].

In the auxiliary methodology of pattern recognition, an accumulation of complicated patterns are depicted by various sub-designs and the syntactic standards with which these sub patterns are connected with one

another. This model is interested with structure and endeavors to perceive a pattern from its standard structure. The accent, which gives the basic characterization of a pattern as far as pattern primitives and their arrangement is termed as PDL (Pattern Description Language). The expanded elucidating force of a language prompts the expanded unpredictability of the sentence structure examination framework.

To perceive finite-state languages (a limited or vast arrangement of strings (sentences) of the images (words) produced by a limited arrangement of principles (the language structure), where every guideline determines the condition of the framework in which it can be connected) is utilized. Comprehensive influence of FSL is less stronger than that of the context-sensitive languages. Context-sensitive languages (formal language that can be characterized by a connection touchy sentence structure (and identically by a non-contracting punctuation). That is one of the four sorts of linguistic uses in the Chomsky progressive system.) are represented by non-deterministic strategies. Determination of sort of syntax for example characterization relies on the primitives and on the language structure's Comprehensive power and evaluation effectiveness [18].

For characterization of patterns, for patterns, chromosome pictures, 2D-science, chemical based structures, verbal words, English figures and unique finger impression designs, and various languages are generally recommended [20]. High magnitude pattern requires high dimensional language structures, for example, web linguistic uses, tree sentence structures, chart punctuations and shape syntaxes for effective depiction [21]-[23].

Stochastic languages, guess and transformational sentence structures are utilized to explain noisy and mishaped designs [26]. This methodology requests extensive preparing sets and expansive computational endeavors at the point when managing uproarious patterns, linguistic use characterizing the fundamental structure of complex patterns that looks to be extremely problematic, making it impossible to characterize, there in such cases, measurable methodology is a decent choice. Acknowledgment Error is the standard to

determine the overall performance. This model is utilized as a part of the application territories like in textured pictures, shape investigation of forms and picture understanding where designs have a distinct structure [27].

2.3 Template Matching Model

Template matching or Format coordinating is least complex and most primitive between each pattern recognition models. It is utilized to decide the resemblance between a pair of pattern, pixels or bends. The pattern to be identified is coordinated with the put away saved template, while accepting that layout can be experienced by rotational or scalar changes.

The effectiveness of this model relies on the saved layouts. Connection capability is used as an acknowledgment ability and is enhanced relying upon the accessible preparing set. The deficiency of this methodology is that it doesn't work proficiently within the sight of mishaped distorted patterns [28].

2.4 Neural Network Based Model

Neural systems are the greatly parallel structures made out of "neuron" like subunits. Neural systems gives proficient result in the field of characterization. Its feature of changing its weight iteratively and learning [10], [29], gives it a side over different strategies for acknowledgment process.

Perceptron is a primitive neuron model. It is a two-layer framework. One of the off possibilities is that, it yields capacity of perceptron in this step, and then it performs classification issues. If there is a chance that it is direct (linear), then it performs regression issues [6]. The most ordinarily utilized group of neural systems for pattern arrangement & classification is the feed forward systems like RBF systems. Distinctive sorts of neural systems are utilized relying on the necessity of the application.

General Regression Neural Network (GRNN) is an exceptionally parallel structure in which, training is from the insight side to the output side [31]. Feed Forward Back-Propagation Neural Network (FFBP-NN) is utilized to execute the non-direct (non-linear) differentiable functionality. The increment in the learning rate in back-proliferation neural system prompts diminish in joining

time [30].

General Regression Neural Network (GRNN) executes productively on noisy information than Back-propagation. FFBP Neural Network is not going to work precisely if the accessible information is sufficiently huge. Then again in GRNN, as the measure of information builds, the mistake gets near towards zero [31]. Kohonen- Networks tend to be for the majority part utilized for information bunching and highlight mapping [14]. Ripley [32] and Anderson et al. [33] expressed the relationship between neural systems and analytical or measurable model of pattern recognition.

The execution of the neural system upgrades after expanding the amount of concealed layers up to a specific degree. The expanded quantity associated with neurons in the covered layer, additionally enhances the execution of the framework. A number of neurons are required to be sufficiently vast to satisfactorily speak to the issue area and sufficiently little to allow the speculation from preparing the information. An exchange off must be kept up to a size of a system and a multifaceted nature came about on account of system size. Rate acknowledgment precision of a neural system might be further improved on the off chance that we utilize 'fuzzy' blend of actuation capacities for neurons of the concealed layer and the yield layer selected as against deciding on different mixes [34].

2.5 Fuzzy Based Model

The significance of fuzzy sets in Pattern Recognition lies in demonstrating the types of instability that can't be completely comprehended by the utilization of probability theory [35], [36]. Kandel declares, "In an exceptionally central manner, the personal connection between the hypothesis of fuzzy sets and a hypothesis of Pattern Recognition and arrangement lies in the way that, most real classes are fuzzy in a universe", Kandel characterized different procedures of the fuzzy pattern recognition. Syntactic methods are used when the pattern looked for is identified with the formal structure of language. Semantic procedures are utilized when the fuzzy segments of information sets are to be created. At

that point, a comparative measure in light of the weighted distance is utilized to acquire a closeness degree between the fuzzy description of obscure shape and reference shape.

2.6 Hybrid Model

In the vast majority of the rising applications, unmistakably a solitary model utilized for grouping doesn't act effectively, so various strategies must be consolidated together offering the result to crossover models. Primitive ways to deal with the configuration, a Pattern Recognition framework which goes for using a best individual classifier have a few disadvantages [37].

It is extremely hard to distinguish the best classifier unless the profound earlier information is accessible nearby [38]. Analytical and Structural models can be joined together to take care of the hybrid issues. In such cases, Statistical methodology is used to perceive design primitives and syntactic methodology is then utilized for the acknowledgment of sub-patterns and pattern itself. Fu [27] gave the approach of assigned grammars, which unifies the analytical and constructive pattern recognition approach.

To improve framework execution, one can utilize an arrangement of individual classifiers and the combiner to settle on a definite conclusion. Tumer and Ghosh [28] tentatively presented the fact that using a linear combiner or demand knowledge combiner reduces the variation of real option limitations around the ideal limit.

Different classifiers can be utilized as a part of a few approaches to improve the framework execution. Every classifier can be prepared in an alternate locality of the highlight space or in other way, every classifier can give a likelihood evaluation and choice can be made after examining the singular results. Techniques using classifier ensemble layout [39], produce an arrangement of the commonly correlative classifiers that accomplish ideal exactness utilizing a fixed decision function. Those strategies which use the blend capacity layout tend to locate an ideal mix of choices from an arrangement of classifiers. To accomplish ideal results, a huge arrangement of mix elements of expanding many-sided

quality, extending from straightforward voting rules through trainable blend capacities is accessible to the fashioner [40],[41].

Conclusion

A near perspective of all the models of pattern recognition has been indicated which shows that for different spaces around these various models or combination of models can be utilized. If generally there must occur an event of noisy patterns, decision of factual model is a decent arrangement. Down to earth significance of basic model relies on recognition of straightforward pattern primitives and their connections spoke to by characterization language. When compared to factual pattern recognition, auxiliary pattern identification is a more up to date zone of evaluation. For complex pattern and applications using a huge number of pattern classes, it is gainful to depict every pattern as far as its segments. An insightful choice with respect to the determination of Pattern linguistic use impact calculations proficiency of the recognition framework. Design primitives and pattern language structure to be used relies on the application prerequisites.

Low reliability of neural systems on earlier information and accessibility of efficient learning algorithms have made the neural systems popular in the field of Pattern Recognition. Even though neural systems and statistical pattern recognition models have diverse standards, a large portion of the neural systems is like factual statistical pattern identification models. To perceive obscure shapes, the fuzzy strategies are great choices. As every model has its own upsides and downsides, along these lines to improve framework execution for complex applications, it is valuable to add two or more recognition models at different phases of the recognition method.

References

- [1]. H.M. Abbas and M.M. Fahmy, (1994). "Neural Networks for Maximum Likelihood Clustering". *Signal Processing*, Vol.36, No.1, pp.111-126.
- [2]. H. Akaike, (1974). "A New Look at Statistical Model Identification". *IEEE Trans. Automatic Control*, Vol.19, pp.716-723.
- [3]. S. Amari, T.P. Chen, and A. Cichocki, (1997). "Stability Analysis of Learning Algorithms for Blind Source Separation". *Neural Networks*, Vol.10, No.8, pp.1,345-1,351.
- [4]. A. Antos, L. Devroye, and L. Györfi, (1999). "Lower Bounds for Bayes Error Estimation". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.21, No.7, pp.643-645.
- [5]. J.C. Bezdek, (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press,.
- [6]. C.M. Bishop, (1995). *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press.
- [7]. C.J.C. Burges, (1998). "A Tutorial on Support Vector Machines for Pattern Recognition". *Data Mining and Knowledge Discovery*, Vol.2, No.2, pp.121-167.
- [8]. B. Cheng and D.M. Titterton, (1994). "Neural Networks: A Review from Statistical Perspective". *Statistical Science*, Vol.9, No.1, pp.2-54.
- [9]. P.A. Chou, (1991). "Optimal Partitioning for Classification and Regression Trees". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.13, No.4, pp.340-354.
- [10]. T.M. Cover, (1992). "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition". *IEEE Trans. Electronic Computers*, Vol.14, pp.326-334.
- [11]. P.A. Devijver and J. Kittler, (1989). *Pattern Recognition: A Statistical Approach*. London: Prentice Hall.
- [12]. L. Devroye, (1988). "Automatic Pattern Recognition: A Study of the Probability of Error". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.10, No.4, pp.530-543.
- [13]. U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, (1999). "Knowledge Discovery and Data Mining: Towards a Unifying Framework". *Proc. Second Int'l Conf. Knowledge Discovery and Data Mining*.
- [14]. Anil K Jain, Robert P.W. Duin and Jianchang Mao, (2000). "Statistical Pattern Recognition: A review". *IEEE Transaction on Pattern Analysis and Machine*

Intelligence, Vol.22, No.1, pp.4-37.

[15]. Scott C Newton, Surya Pemmaraju and Sunanda Mitra, (1992). "Adaptive Fuzzy Leader Clustering of Complex Data Sets in Pattern recognition". *IEEE Transactions on Neural Network*, Vol.3, No.5, pp.794-800.

[16]. Jenn- Hwai Yang and Miin Shen Yang, (2005). "A Control Chart Pattern Recognition System Using Statistical Correlation Coefficient Method". *Computers and Industrial Engineering*, Vol.48, pp.205-221.

[17]. Jun Hai Zhai, (2006). "An Overview of Pattern Classification Methodologies". *Proceedings of 5th International Conference on Machine Learning and Cybernetics*, pp.3222-3227.

[18]. Liu, J., Sun, J. And Wang, S. (2006). "Pattern Recognition: An overview". *International Journal of Computer Science and Network Security (IJCSNS)*, Vol.6, No.6.

[19]. Vapnik, V., (2010). *The Nature of Statistical Learning Theory*. Springer.

[20]. Robert Jain, and Mao, (2000). "Statistical pattern recognition: A Review". *IEEE Transaction on Machine Learning and Pattern Analysis*, Vol.22, No.1.

[21]. Wikipedia. *Introduction to pattern recognition*.

[22]. Thomas Minka. *A Statistical Learning/Pattern Recognition*.

[23]. Zheng, He, (2011). *Classification Technique in Pattern Recognition*, University of Technology, Sydney, ISBN 80-903100-8-7, WSCG.

[24]. Seema Asht, and Rajeshwar Dass, (2012). "Pattern Recognition Techniques: A Review". *International Journal of Computer Science and Telecommunications*.

[25]. M. Parasher, S. Sharma, A.K Sharma, and J.P Gupta, (2011). "Anatomy on Pattern Recognition". *Indian Journal of Computer Science and Engineering (IJCSE)*, Vol.2, No.3.

[26]. Amin Fazel and Shantnu Chakrabartty, (1996). "An Overview of Statistical Pattern Recognition Techniques for Speaker Verification". *IEEE Circuits and Systems*, pp.61-81.

[27]. K.S. Fu, (1982). *Syntactic Pattern Recognition and Applications*. Englewood Cliffs, N.J.: Prentice-Hall.

[28]. K. Tumer and J. Ghosh, (1996). "Analysis of Decision Boundaries in Linearly Combined Neural Classifiers". *Pattern Recognition*, Vol.29, pp.341-348.

[29]. R.O.Duda and P.E.Hart, (1973). *Pattern Classification and Scene, Analysis*, New York: John Wiley & sons.

[30]. Majida Ali Abed, Ahmad Nasser Ismail and ZubadiMatiz Hazi, (June, 2010). "Pattern recognition Using Genetic Algorithm". *International Journal of Computer and Electrical Engineering*, Vol.2, No.3.

[31]. Mohammad S. Alam, and Mohammad A. Karim, (2004). "Advances in Pattern Recognition Algorithms, Architectures and Devices". *Optical Engineering*, Vol.43, No.8.

[32]. B. Ripley, (1993). *Statistical Aspects of Neural Networks, Networks on Chaos: Statistical and Probabilistic Aspects*. U. Bornndorff-Nielsen, J.Jensen, and W. Kendal, Eds., Chapman and Hall.

[33]. J. Anderson, A. Pellionisz, and E. Rosenfeld, (1990). *Neuro Computing 2: Directions for Research*. Cambridge Mass.: MIT Press.

[34]. T. Pavlidis and F. Ali, (1975). "Computer Recognition of Handwritten Numerals by Polygonal Approximations". *IEEE Trans. Syst., Man, Cybern.* Vol. SMC-5.

[35]. Anupam Joshi, Narendram Ramakrishman, Elias N. Houstis and John R. Rice, (1997). "On Neurobiological, Neuro-Fuzzy, Machine Learning And Statistical Pattern Recognition Techniques". *IEEE Trans. on Neural Networks*, Vol.8, No.1.

[36]. J.C. Bezdek, (1981). *Pattern Recognition with Fuzzy Objective Function Algorithm*, New York: Plenum Press.

[37]. R. O. Duda, P. E. Hart and D. G. Stork, (2000). *Pattern Classification*, John Wiley & Sons.

[38]. B. Ripley, (1996). *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge.

[39]. Sergios Theodoridis, and Konstantinos Koutroumbas, (1982). "Pattern recognition". *Pattern Recognition*, Elsevier, USA.

[40]. Anil k Jain, and Robert P.W Duin, (2004). *Introduction To Pattern Recognition*. The Oxford Companion to the Mind, Second Edition, Oxford University Press, Oxford, UK,

pp.698-703.

[41]. S.P. Shinde and V.P. Deshmukh, (2011).
"Implementation of Pattern Recognition Techniques and

Overview of its Applications in various areas of Artificial Intelligence". *International Journal of Advances in Engineering & Technology*, Vol.1, No.4, pp.127-137.

ABOUT THE AUTHORS

Navneet Kumar Kashyap is pursuing his M. Tech. at Govind Ballabh Pant University Agriculture & Technology, Pantnagar, Uttarakhand, India in the Department of Information Technology. He received his B.Tech. degree in Computer Science & Engineering from SLSET Group of institutions, Affiliated to Uttarakhand Technical University, Dehradun, India in 2013. His interest includes Big-Data, NoSQL and Web Technologies & Application.



Binay Kumar Pandey is currently working as an Assistant Professor in the Department of Information Technology, College of Technology, GB Pant University of Agriculture & Technology, India. His areas of interest includes High Performance Computing, Bio-informatics, and Cloud-Computing.



Dr. Hardwari Lal Mandoria is currently working as a Professor & Head In the Department of Information Technology, College of Technology, GB Pant University of Agriculture & Technology, Pantnagar, India. His areas of Interest are Computer Networks, Network Security Wireless Communication, Mobile Computing, Information Security, and information Communication Technology.



Ashok Kumar is currently working as an Assistant Professor in the Department of Information Technology, College of Technology, GB pant University of Agriculture & Technology, India. His area of interest includes Software Engineering, Software Testing and Software Analytics.

