

CHAPTER 55
TECHNOLOGY
COMPUTER ENGINEERING

Doctoral Theses

01. BENIWAL (Rohit)
Social and Semantic Web Capabilities for Improving Software Development Activities.
Supervisors :Dr. M.P. S. Bhatia and Dr. Akshi Kumar
Th 24740

Abstract
(Verified)

Software development is a multifaceted task with new challenges being imposed frequently. The problem manifolds owing to the overwhelming demands of the customers, numerous participants, shorter time frames, different geographical and virtual locations of software development teams, information overload & other related issues. Sharing and reusing already available information saves the effort in development and maintenance of software systems. Engaging all stakeholders in development process produces a desirable software. Taking crowdsourcing help further reduces the development effort. However, there are major operative challenges that foster the need to explore strategic and supporting technologies to tackle challenges such as efficiently reusing and sharing of software development knowledge, effectual involvement of all stakeholders particularly users in the software development process, and efficacious engagement of crowdsourcing in the software development process. In this regard, recent studies have shown the collaboration among research fields of (1) Social Web and Software Engineering and (2) Semantic Web and Software Engineering, which clearly demonstrate the prospective benefits of applying Social and Semantic Web capabilities for improving software development activities. This collaboration addresses various software development challenges and provides opportunities to resolve them. As far as our work is concerned, we also used the Social and Semantic Web capabilities for improving software development activities by providing the solution to tackle few software development challenges.

Contents

1. Introduction 2. Literature review 3. Problem statement 4. Overview of work 5. Conclusion and future work 6. References and List of publications.

02. GUPTA (Shikha)
Cryptographic Techniques for Enhancing the Data Security.
Supervisor :Prof. Satbir Jain
Th 24732

Abstract
(*Verified*)

This thesis focuses on various encryption techniques to achieve the goal of security services by considering numerical, alphanumeric and alphabetical data. Firstly, the thesis starts with an overview of cryptography which defines the various symmetric encryption ciphers used in cryptography such as AES, DES, OTP, etc. These ciphers are suitable for encrypting the data but are unsuitable for encrypting sensitive data items as they lead to change in the structure of the encrypted data items to be stored in the database. The objective of this thesis is to design new and efficient cryptographic algorithms for maintaining the structure and improving the performance of sensitive data items of numeric and alphanumeric format and then storing the encrypted data items back to the database in their original format. Advanced encryption standard (AES) is used as a base cipher, which comprises mathematical operations like addition, subtraction, multiplication and conversion methods. The second part of the thesis defines a new encryption technique in which AES 128-bit encryption cipher is used to protect the data and which can be moved up to 256-bit encryption if needed along with the enhanced One-time pad (OTP) encryption cipher. A method for generating random keys is also proposed known as a Secure key generator (SKG) to add more security to the whole system. This part of the thesis carries a detailed study of AES and OTP encryption cipher as well as compares the results of enhanced OTP cipher with existing encryption ciphers. Finally, in the last part of the thesis, Neural cryptography and Artificial neural network are used for generating keys and how these keys are helpful in performing encryption. The process of mutual learning is also used for exchanging keys by using a single layer network known as Tree Parity Machine (TPM).

Contents

1. Introduction 2. Literature survey 3. Length preserving symmetric encryption based on block cipher mode of operation 4. A Data encryption scheme to encrypt numeric data in database by using advanced encryption standard block cipher 4. DWSA: a secure data warehouse architecture for encrypting data using AES and OTP encryption 7. An encryption approach to improve the security and performance of data by integrating AES with modified one time pad(OTP) 8. Mutual learning in tree parity machines using cuckoo search algorithm for secure public key exchange 9. Conclusion and future work 10. References 11. List of publication.

03 KAUR (Manpreet)

New Methods for Enhancing the Quality of Document Summarization Using Textual Entailment.

Supervisor : Dr. Anand Gupta

Th 24900

Abstract
(*Verified*)

Digital revolution has led to the problem of information overload in society. Automatic Text summarization is a viable solution to this problem of excessive accumulation of information. However the goal of generating highly informative and coherent summaries is yet not achieved. Therefore, we propose new text summarization methods to improve the quality of a summary. The Literature survey motivates us to employ a recent approach of Textual entailment (TE) in the proposed methods. It is an established indicator of semantic relationships between text units. It can be used to measure sentence connectedness. The foremost method of single document summarization is based on employing TE to bring related sentences under clusters representing various themes. The method is

termed as Analog Textual Entailment and Spectral Clustering (ATESC). It is extended to propose new methods for multi-document summarization also where SC handles the issue of redundancy well. The experimental results show improvement in informativeness of a summary but they still lack coherence. Therefore, we propose to use a graph-based method which cast the text summarization task as an optimization problem. It is solved using weighted Minimum Vertex Cover algorithm (wMVC), a graph-based algorithm. The method is termed as Textual Entailment based Minimum Vertex Cover (TEMVC). However, the performance of the proposed methods largely depend on the accuracy of TE recognition method. Thus, we hypothesize to use a better approach of entailment, named as Partial Textual Entailment in the next summarization method. The method formulates the single document summarization task as a set based optimization problem which is solved using weighted Minimum Set Cover (wMSC) algorithm. The method is termed as Partial Entailment based Minimum Set Cover (PE-MSC). The experimental results show that the proposed methods yield competitive results.

Contents

1. Introduction 2. State of the art 3. Summarization using analog textual entailment based spectral clustering 4. Summarization using entailment based graphs 5. Summarization using entailment based sets 6. Conclusion and future directions. Bibliography . List of publications and Appendix.

04. KAUR (Preeti)

Qualitative and Quantitative Assessment of Requirements Engineering Artifacts.

Supervisors : Prof. Sangeeta Sabharwal and Prof. Ritu Sibal
Th 24900

Abstract (Not Verified)

In today's world, there is growing demand for solving complex problems through software systems. A crucial factor in the development of good quality and cost-effective software is effective requirements understanding. Any defects introduced during the requirements engineering phase, increases complexity and cost of fixing them later in the development cycle. A deeper understanding of the requirements of the problem domain is realized by making analysis models. These models form the set of requirements engineering artifacts. The main aim of this research work is to assess the requirements engineering artifacts based on their structural properties through qualitative and quantitative techniques. A complexity metric based on UML use case diagram has been derived from the dependency relations that exist in it. The relationship between the structural complexity and the external attribute understandability has been studied. This proposed approach has been empirically and theoretically validated. A technique widely used for safety, risk and reliability analysis called Fault Tree Analysis (FTA) has been applied to assess the reliability during early software development. FTA has been used with use case diagram and activity diagram to derive and validate a reliability metric. Additionally, the application of web ranking algorithms on the UML use case diagram is demonstrated. To extend our study regarding the artifacts of requirements phase, we also explored the paradigm of Agile Software Development. User story conceptual model has been proposed and used to derive a systematic approach for computing system complexity in this paradigm.

Contents

1 Introduction 2. Literature review 3. Empirical and theoretical validation of a use case diagram complexity metric 4. Deriving and validating fault metric for

object oriented system using use case diagram 5. Applying page rank and HITS algorithm to identify key use case 6. A user story based approach to measure system complexity in Agile software development 7. Conclusion and future work 8. Bibliography.

05. KUCHHAL (Pragaya)
Efficient and Reliable Routing Techniques for Data Forwarding in Opportunistic Networks.
 Supervisor :Prof. Satbir Jain
Th 24731

Abstract
(Verified)

Opportunistic Network (OppNets) is one of the leading domain of wireless communication. It works on the principle of store-carry-forward. Whenever a nodes doesn't find the suitable node to further transmit the message the node store the message in its buffer. As and when finds the suitable node forwards the message to it. Routing is very crucial aspect in this network for transmitting the data packets successfully to its desired destination. An ideal routing technique is one which is able to deliver the message to its destination with minimum delay and cost. In this context, few of the optimized and efficient routing schemes have been designed. In first approach "A*OR" the A* search algorithm is implemented in locating the best and worthy nodes for transmitting the data packets in the network. The second approach Pareto set based Optimized Routing (PORON) focuses on optimized routing through multi-objectives based approach. This approach involves designing of new weighted function to select the best next optimized set of nodes from the combination of parameters, namely; maximizing energy, minimum signal-to-noise plus-interference-ratio (SINR) and maximizing contact duration. Thirdly, performance evaluation of A*OR is implemented through different mobility models. Nextly, energy proficient scheme "EA*OR" is designed. This scheme considers the residual power/energy of the current and neighboring node. On the basis of the efficient energy the nodes are selected and message is transmitted to them. Lastly a security measurement is dealt in OppNets. Under this framework black hole nodes are identified and isolated based on the forwarding ratio. The analysis of the algorithms was performed on the Opportunistic Network Environment (ONE) simulator. ONE is one of the efficient and best simulation tool for opportunistic networks.

Contents

1. Introduction 2. Literature review 3. A * search based next hop selection for routing in opportunistic networks 4. Pareto set based optimized routing in opportunistic network 5. Performance analysis of A* OR on mobility modes 6. Energy proficient A* perception in infrastructure less opportunistic networks: EA*OR 7. Blackhole mitigation in A* search based hop selection routing in opportunistic networks 8. Conclusion and future directions 9. List of publications and References.

06. MITTAL (Ruchi)
Design and Analysis of Clustering Methods for Social Networks.
 Supervisors :Dr. M.P. S. Bhatia and Akshi Kumar
Th 24738

Abstract
(*Verified*)

The cloud Due to advancements in technology, the usage of social media has increased at a rapid rate. There are many social media applications such as Twitter, Facebook, WhatsApp, and many more where people are connected and form a network called a social network. Since social interaction among users grown at an exponential rate, various real-life social networks exhibit multiple types of communication among entities. Thus, this arrangement is known as multi-layer networks. Measuring centralities, structural holes, and identifying influential users in such a system are challenging tasks. Analyzing and identifying clusters or communities have rapidly grown as one of the prominent topics in social networks. A community is a close-knitted set of nodes which plays an indispensable useful part in the primary system. In this thesis, we marked the following issues: Propose a similarity-based clustering algorithm for finding the clusters or communities from multi-layer social networks. Design new methods for measuring the closeness centrality and structural holes in multi-layer networks. We also collaborate with nature-inspired algorithms with social network techniques to get the improvised techniques of closeness centrality and structural holes. Propose methodologies for finding the most stressed and bottleneck entities of the system, which have a significant impact on the network. Address the problem of anomaly detection, identification of suspicious users from a single layer, and multi-layer social networks. Present a diversified ensemble approach for finding the most influential users of the system in various application domains.

Contents

1. Introduction 2. Literature review 3. Similarity based clustering algorithm for multi layer social networks 4. Measuring closeness centrality & structural holes in multi layer social networks 5. Discovering bottleneck entities and stress entities in multi layer social network 6. Anomaly detection in multiplex social networks 7. Detection of suspicious or Un-trusted users in cryptocurrency financial trading applications 8. Classifying the influential individuals in multi layer social networks 9. Conclusion and future work . List of publication and References.

07. POONAM RANI
A Novel Approach to Social Network Analysis Using Soft Computing Techniques.
 Supervisors :Dr. M.P. S. Bhatia and Dr. Devendra K. Tayal
Th 24739

Abstract
(*Not Verified*)

Web2.0, with its tools, assists people to easily collaborate, share and maintain contact globally. The social network is one such tool that plays a vital role in our lives. Social networks are the most dynamic, powerful and all-purpose tools for people of all classes. They use structured, unstructured, multilingual, audio, video and all other types of data efficiently. Most of the time, we need to find their role, advantage, effectiveness, and outcomes in our day-to-day lives. This further helps us to employ these social networks for a variety of applications. A comparison of social networks helps us to investigate which of them is better in terms of connectivity factors. The most important point is that these different types of social networks are highly uncertain. Their parameters are uncertain, variable and possess dynamic properties. So, we use soft computing techniques to propose direct and indirect techniques to compare social networks. In direct techniques, we

propose an algorithm that directly compares social networks on a fixed set of network parameters or metrics at three levels - network, group, and node level. For this, we propose three algorithms by using a quantitative approach, a qualitative approach and a mixed approach based on user experiences towards social networks. We propose a three-dimensional conical model for the mixed approach. On the other hand, we propose three algorithms for indirect techniques. In this, we propose functions and approaches that indirectly compare social networks. In this, we predict the potential relationship of the social network that can further be used for the comparison of social networks. This prediction parameter provides a global indicator of the social network. So, we propose an astute socio-centric methodology for aggregating the potential relationship in the network using a fuzzy graph social network model. In this, we propose three fuzzy definitions.

Contents

1. Introduction: Social network and social networks analysis 2. Literature review 3. Direct methodologies 4. Indirect methodologies 5. Conclusion and future scope. References and List of publications.

08. RAI (Sunny)
Computational Metaphor Processing For English Language.
 Supervisor : Prof. Shampa Chakraverty
Th 24733

Contents

1. Introduction 2. Literature Survey 3. Supervised detection using conditional random fields 4. Impact of context on metaphor detection 5. Understanding metaphors using emotions 6. Modeling imprecision and approximation in metaphor detection 7. Conclusion 8. Classification techniques 9. Bibliography .

09. SARASWAT (Mala)
Investing Qualitative Behavior in Cross Domain Recommender Systems .
 Supervisor : Prof. Shampa Chakraverty
Th 24737

Contents

1. Introduction 2. Survey of literature 3. Modeling keyword similarity and genre correlation for cross domain recommendation 4. Enriching topic coherence on review for cross recommendation 5. An emotion driven recommender system 6. Review based emotion profile for cross domain recommendations 7. Conclusion and future scope 8. References and List of publication.

10. SARNA (Geetika)
A Novel Approach to Model Anomalous Behavior of User in Social Network.
 Supervisors : Dr. M.P. S. Bhatia and Akshi Kumar
Th 24741

*Abstract
 (Verified)*

Social Network is a form of social media used for interaction, education, information sharing and entertainment purposes. Since, with the advancement of technologies and connectivity, more people have been joined and spend most of their time on social network. The

excessive use of online social network may give rise to various types of anomalous behavior such as phishing, identity theft, cyberbullying and many more. Cyberbullying is one of the types of anomalous behavior and also has negative repercussions which need to be handled. Therefore, in this thesis we considered cyberbullying as anomalous behavior. Cyberbullying is to post harmful, harassing material or to engage in other forms of social aggression using digital technologies that may results in several emotional, psychological and physical problems, poor academic record and also results in suicidal ideation. The thesis starts with the analysis of the relation between cyberaddiction and cyberbullying followed by the detection of cyberbullying for which two approaches have been used: First approach identify cyberbullying on the basis of the severity of the bullying in messages and second approach identify the cyberbullying on the basis of direct and indirect bullying behaviour . Main limitation of above discussed two approaches is that it depends upon the existing lexicon of bad words. So, the updation of lexicon with new bad words has also been covered in thesis. The next contribution of this thesis is to identify the different categories of Cyberbullying that would be helpful in the identification of the harmful sector to avoid the incidents. The next contribution of this thesis is to find the fake profiles responsible for cyberbullying. Two approaches have been used: (1) Identification of Fake Profiles using Entropy and (2) Identification of Suspicious Patterns using Zipf's Law in Social Network. In this thesis, we have used dataset of Twitter..

Contents

1. Introduction 2. Literature Survey 3. Cyberbullying: a risk of cyber addiction 4. Identification of cyberbullying 5. Identification of new keywords responsible for cyberbullying 6. Analysis of different Categories of cyberbullying 7. Identification of fake responsible for cyberbullying 8. Conclusion and future work 8. References. Appendices and List of publications.

11. SHARMA (Deepak)
Designing Algorithms for Trends Analysis in Research.
 Supervisors : Prof. Bijendra Kumar and Prof. Satish Chand
Th 24744

Abstract (Verified)

Trend Analysis is a practice of gathering information and endeavoring to spot a pattern or trend in the information. Due to the vast number of scientific papers publishes very rapidly, it becomes tedious for the researchers to stay up to date with new research themes. These themes are naturally arising from the statistical characteristic of the documents, and as such no prior annotation or labeling is required. In order to understand about the topic modeling, a survey presented on topic modeling techniques since its inception from LSA to the latest topic models in deep learning techniques for generating topics from collection of documents and to understand the trend analysis in different research areas. Trend analysis in machine learning area is unexplored and there is a need to identify trends. We have collected the titles and abstracts research data from journals published in machine learning. The corpus has prepared by collecting 23,365 articles from the time span of 50 years from 1968~2017. The objective of this thesis is to design algorithms and framework to identify trends in machine learning research. The first work focuses on identify the trends in machine learning research using text mining. The second work focuses on to systematically examine the literature of machine learning to identify and analyze the research trends using topic models and Mann-Kendall test. The third work focuses on a topic network analysis approach which integrates topic modeling and social network analysis. The fourth work focuses on research topic over time approach using topic coherence model using LDA. The fifth work focuses on to uncover the researchers in machine learning using author-topic model (ATM). Finally, the sixth work focuses to uncover research topics in machine learning communities with

scientific collaboration network (SCN). The experimental results demonstrate that our simple technique is efficient and feasible.

Contents

1. Introduction 2. Literature survey 3. Dataset Description 4. Trends analysis in research using text mining 5. Trends analysis in research with topic models and mann kendall test 6. Trends analysis in research using topic network analysis 7. Research topics over time: a trend analysis using topic coherence model with LDA 8. Identify and recommending researches with author topic model 9. Uncovering research trend and topics of communities in research 10. Conclusion and future work 11. List of publications and References.

12. SHARMA (Ruchi)

Assessment Frameworks in Software Security and Reliability Engineering.

Supervisors :Prof. Ritu Sibal and Prof. Sangeeta Sabharwal

Th 24899

Abstract (Verified)

The research study in this thesis has been conducted in two main domains viz. Vulnerability Discovery Modelling and Software Vulnerability Prioritization. Software Vulnerability Discovery models gives a quantitative approach towards the issue of vulnerability management. After a vulnerability is discovered, the management faces an important issue of deciding which vulnerability is to be fixed first. For this testing team have to prioritize the vulnerabilities so that immediate action can be taken to fix vulnerability possessing higher risk. In the present research work, we propose to address software security by addressing software vulnerabilities. In this thesis we have discussed the shortcomings of existing vulnerability discovery models (VDMs) and proposed some new models by considering data independence, change point and multilevel severity in the software vulnerabilities. We have used Statistical Package for Social Sciences (SPSS) for parameter estimation of the proposed models. The applicability of the discussed models has been shown by validating it on actual software vulnerability data sets. We have also proposed new prioritization techniques so as to assist the testing team in the process of vulnerability fixation once the vulnerabilities have been detected. In a real software program, one vulnerability can trigger or inhibit the existence of another vulnerability. Therefore interdependence among vulnerabilities can increase or decrease the overall risk posed to the system. So, under this consideration, we have proposed a multi criteria decision making (MCDM) technique for vulnerability prioritization. We have further conducted a comparative study for software vulnerability prioritization using two MCDM techniques viz. TOPSIS and VIKOR on real life software vulnerability datasets. Finally we have proposed a method to eliminate the need for generating a severity score for software vulnerabilities by proposing a new technique for prioritization using the description of a vulnerability using machine learning.

Contents

1. Introduction 2. Literature survey 3. A new approach to vulnerability discovery modelling 4. Multilevel vulnerability discovery modelling and risk assessment 5. Software vulnerability prioritization as a decision making problem 6. A new outlook towards software vulnerability prioritization 7. Conclusion and future research directions. List of publications and References.

13. SINGH (Yash Veer)
Design and Analysis of Prioritization Techniques for Software Requirements.
 Supervisor : Prof. Bijendra Kumar and Prof. Satish Chand
Th 24742

Abstract
 (Not Verified)

The challenge to the software industry communities is to satisfy the customer's need, their multiple criteria and possibly exceed their expectations in an economic, rapid and profitable manner. If a system has thousands of requirements in the form of linguistic variables then it is very difficult for a project manager to prioritize such a large number of requirements. Multiple stakeholders provide different opinions; so many conflicts arise in the process of requirements prioritization. It makes the requirements prioritization process unreliable, more time consuming that influences the overall software product's quality, ordering, and software release planning. So the main focus of requirements prioritization techniques is to evaluate the ranking of requirements as per their importance of order and to implement the highest priority requirements first with respects to several aspects like, quality, cost, risk, and software release time. From the literature survey, it has been found that none of the techniques of requirements prioritization works well for ranking of requirements and they suffer from a number of limitations like complexity, uncertainty, ambiguity and negative degree of membership function etc. The motivation of this thesis is to design and analyze of efficient requirements prioritization techniques to rank the requirements as per their importance of order in the software development process. Our proposed approaches resolve the classical gaps and to meet the client fulfillment of decision making in real applications. In this thesis we have proposed five requirements prioritization techniques using the following approaches: (i) logarithmic fuzzy trapezoidal approach (LFTA), (ii) a hybrid approach using artificial neural networks (ANNs) and fuzzy analytical hierarchy process (AHP) model, (iii) a hybrid approach using logarithmic fuzzy preference programming (LFPP) and ANN, (iv) a hybrid approach using fuzzy AHP with particle swarm optimization (PSO), and (v) a hybrid approach using LFTA with ANN.

Contents

1. Introduction 2. Literature review 3. Requirements prioritization using LFTA 4. Requirements prioritization using ANN fuzzy AHP model 5. Requirements prioritization using Hybrid LFPP and ANN 6. Requirements prioritization using hybrid approach fuzzy AHP with PSO 7. Requirements prioritization using hybrid approach LFTA and ANN 8. Conclusion and future scope 9. References and List of publication.

14. SINGH (Deepti)
Designing of Authentication and Message Security Schemes in Wireless Sensor in Wireless Sensor Network.
 Supervisors : Prof. Bijendra Kumar and Prof. Satish Chand
Th 24898

Abstract
 (Verified)

The wireless sensor networks (WSNs) are being used in numerous applications such as disaster management, smart building monitoring, intelligent home appliances, etc. A WSN consists of large number of sensor nodes that sense the environment and send the collected data to the base station for further processing. In order to avoid unauthorised access, user authentication is essential for accessing the information from WSNs. The

mechanism for accessing data securely from the WSNs can be user authentication, including mutual authentication. For secure communication, user, base station and sensor node mutually authenticate each other. Since the sensors in WSNs have limited resources in terms of processing power, memory, battery, the traditional security mechanism cannot be applied. Hence, designing of user authentication schemes in WSNs is still a challenging task. In this thesis, we have discussed authentication schemes for WSNs using various approaches. The first work uses MAC function to encrypt the message with the help of a key that is more secure. This scheme also uses symmetric encryption and decryption operations for user authentication. The second work contains two schemes based on Markov chain and biometric. We also propose a secure IoT based mutual authentication scheme that uses elliptic curve cryptography for healthcare applications. In another scheme our work is based on Rabin cryptosystem and Blum Blum Shub pseudo-random generator algorithm. In a biometric-based scheme, digital watermarking is used for the protection of biometric data. The proposed schemes are verified using AVISPA simulation tool to prove that they are resistant to active and passive attacks. The BAN logic has been used to provide mutual authentication and a random oracle model for formal security analysis. Further, security analysis has been done to show that schemes are resistant to various attacks. The proposed schemes perform better than the existing schemes in contemporary literature.

Contents

1. Introduction 2. Literature Survey 3. SMAC-AS MAC and based authentication scheme for wireless sensor networks 4. Markov and biometric based authentication schemes for wireless sensor networks 5. RCBE-AS: Rabin cryptosystem based authentication scheme for wireless sensor networks 6. ECC based mutual authentication scheme for wireless sensor network 7. Conclusion and future scope. List of publications and References.

15. SONI (Rituraj)
To Develop New Methods for Extraction of Text Regions in Natural Scene Images.
 Supervisor :Prof. Bijendra Kumar and Prof. Satish Chand
 Th 24734

Abstract (Verified)

The natural scene images are the pictures clicked by digital or mobile cameras. The text region extraction from these images plays an essential role in finding the textual information in natural scene images. This problem is challenging due to variation in fonts, size, color, different backgrounds, etc. Such textual content assists in domains like aid to visually impaired people, scene understanding, navigation, etc. The methods for extraction of text regions in natural scene images are categorized into edge, texture, connected-components, stroke-based, and hybrid methods. The selection of the text-specific features plays a vital role in discriminating between text and non-text. The choice of classifiers governs the classification of text and non-text elements. The text regions undergo the clustering process to form words. The rectangular boxes are created around the detected text regions. The performance of the method is based on the precision, recall, and f-measure calculated using the ratio of the area covered by the detected rectangles and ground truth available in datasets. These major contributions of thesis are: a) A Fast Edge Preserving Smoothing MSER algorithm is proposed that solves the problem of the interconnection of characters in blurred images. b) A novel method for selecting an optimal feature set from the pool of different features with the help of the Weka tool is proposed. c) A novel method for selecting an optimal classifier from a set of classifiers using the Weka tool is proposed. d) A study regarding the variation of delta (Δ) parameter of MSER & determine its optimal value for

different versions on MSERs is proposed. The training is accomplished on the ICDAR 2013 training dataset, & experiments for testing our method are carried out on ICDAR 2003, 2011, 2013 datasets. The proposed method achieves improved precision, recall, & f-measure on the ICDAR benchmark dataset.

Contents

1. Introduction 2. Literature review 3. Fast edge preserving & smoothing MSER 4. Feature extraction & optimal feature selection 5. Optimal classifier selection 6. Formulation & evolution of text awareness model (TAM): results & discussions 7. Conclusion and future work 8. List of publications. References and Appendices.

16. SRISHTI
Sentiment and Affect Analysis on Social Media.
 Supervisor : Prof. Shampa Chakraverty
Th 24736

Abstract (Not Verified)

Sentiment Analysis is the newest buzzword these days owing to its immense potential to provide real-time feedback. It is a growing field and of interest to researchers in the domains of Computer Science, Social Sciences and Marketing alike. It is, however, still in its nascent stages and improvements are required on many fronts. The objective of this thesis is to investigate the challenges in the field of Sentiment and Affect Analysis on Social Media. We began our research by conducting an extensive survey of the rich body of related research present in literature to assess the strengths and limitations of existing approaches in Sentiment and Affect Analysis. Our investigation revealed that purely Lexicon based and purely Machine Learning based approaches are the ones that are primarily used for Sentiment Analysis and have now attained a point of saturation. These approaches take into account the words in a piece of text and using their polarities as given by a lexicon or some previous training data, determine a cumulative positive or negative sentiment score or polarity for the piece of text in question. Most individuals, organizations, businesses and organizations use tools for Sentiment Analysis loosely derived from these approaches. However, there are still numerous unanswered questions in this field. Previous researchers have not delved on uncovering emotions from text particularly the emotions reflected by emoticons and slangs used in web jargon. Also, the interrelationships between the words in a piece of text and how these context dynamics affect the overall document sentiment is largely unexplored. Moreover, extracting sentiments from text where context changes abruptly continues to remain a challenge. Furthermore, little to no attention has been given on discovering the latent or hidden aspects while considering sentiments of a product. In this dissertation, we approached these issues in a systematic manner.

Contents

1. Introduction 2. Survey of literature 3. Emotion location mapping and analysis using Twitter 4. A context based algorithm for sentiment analysis 5. Tracking context switches I text documents and its application to sentiment analysis 6. Multi aspect sentiment analysis using domain ontologies 7. Conclusion. References. Appendices. List of publications.

17. VEENU

Some Investigation into Training of Feedforward Artificial Neural Networks (FFANN's) and Their Applications.

Supervisor : Dr. M.P.S. Bhatia

Th 24743*Abstract
(Verified)*

Many people are developing interest in soft computing techniques. One of soft computing technique is use of Neural Networks. It has many parameters that could be studied and altered accordingly to our application needs. ANN has number of applications like pattern recognition, function approximation, filtering, control applications. Many models of ANN and training & learning paradigms exists. They have non-recurrent architecture, use supervised training, universal approximation results, intuitive appeal and simple implementation. FFANN using sigmoidal activation function is most widely used for solving number of problem. The training algorithm for FFANN uses arbitrary choices for the parameters of the network. Some of these are The network architecture - considering input, output and hidden nodes. The network weights - considering various weight initialization techniques Algorithm optimization - considering the various training algorithms i.e. supervised and unsupervised algorithms. The activation functions like sigmoidal, tangent or hyperbolic tangent. During literature survey we found that lot of work could be done in this area. Objective of learning is to estimate & learn unknown function mapping from the set of known samples. To tackle any practical problem, one of the main steps used is weight initialization in solving any problem. Many weight initialization methods suggested in various paper use back-propagation algorithm for training the FFANN. . The learning speed of the multilayer FFANN depends on the initial values of its weight. The weight initialization influences the convergence speed and convergence time vary depending on the initialization of weights. The probability of successful convergence also depends on the scheme used for weight initialization. Several papers use random weight initialization. We have statistically analysed the weight initialization method and then proposed weight initialization methods. For comparison four weight initialization methods are considered namely zero, random, Nguyen-Widrow and two proposed weight initialization methods.

Contents

1. Introduction 2. Literature review 3. Feedforward artificial neural networks (FFANN) 4. Weight initializing 5. Experiments 6. Result and observations 7. Conclusion and future work 8. List of publication and References.