

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

LEARNING AND INFORMATION RESOURCE CENTRE

**BIBLIOGRAPHY OF M.E. -
COMPUTERS
DISSERTATIONS**

[1] PR1166 PRIVATE CLOUD SECURITY: SECURED USER
AUTHENTICATION WITH ACCESS CONTROL AND ASSURED
DELETION

By GAJRA, NIKHIL TRIKAM

2014

005.82 GAJ

Project Guide MR.SHAMSUDDIN KHAN

Abstract

Cloud computing is used to store data remotely so that users can take advantage of the on-demand applications and services from a shared pool of configurable computing resources, without the burden of data storage, maintenance and costs. However, since the data is stored by third party cloud providers, it is necessary to provide security guarantees not only on authentication techniques used but also on various files stored over the cloud. This research focuses on techniques like MAES and Blowfish for encryption. A combination of both the techniques i.e. Hybrid technique is also studied and implemented and the results are compared. The structure of S-Box is modified in AES and attempts are made to improve the effectiveness of AES. Diffie-Hellman is used for key generation and key exchange. From the results obtained, the Hybrid technique is more effective in comparison to AES and Blowfish in terms of strength, throughput and time for encryption.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[2] PR1167 IMPROVING LEARNING MANAGEMENT SYSTEM USING DATA MINING

By SOLASKAR, SUPRIYA LAXMAN

2014

005.74 SOL

Project Guide DR A.K.SEN & MRS ANURADHA

Abstract

In the last decade, the effect of internet usage in education has increased tremendously and student's learning skills seems to have improved through these emerging technologies. The use of the Learning Management System (LMS) emerges as a great opportunity to improve and complement teaching and learning, by encouraging students to perform different activities. LMS caters to the needs and usability of both the types of users the faculty/teachers and the students. On one hand it supports teachers in creating, administering and managing online courses. On the other hand, LMS treats all students/learners equally, regardless of their personal abilities to study. This proposed work helps in identifying the different learning styles of students. For this purpose a fuzzy inference system (FIS) is used which recommends a learning material based on student's learning styles and teaching evaluation. The goal is to enhance an existing Learning Management System by delivering course material that best fits student's learning styles.

[3] PR1168 SEMANTIC-BASED DATA MINING AND DATA PRE-PROCESSING FOR BIG DATA

By JAGDALE, ASHISH

2014

005.74 JAG

Project Guide MR.SHAMSUDDIN KHAN

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Big Data is a term which is used to describe massive amount of data generating from digital sources or the internet usually characterized by 3 V's i.e. Volume, Velocity and Variety. From the past few years, data is exponentially growing due to the use of connected devices such as smartphones, tablets, laptops and desktop computers. Moreover E-commerce which is also known as an online market, internet services and social networking sites are generating tremendous user data in the form of documents, emails and web pages. This generated data volume is so vast and overwhelming which makes it complex to process and analyze using traditional software systems and hence consuming more time. Proposed system presents an algorithm to extract user activities such as different types of files and directories accessed from the user computer machine and another algorithm based on Apache's Hadoop Distributed File System known as HDFS framework using MapReduce functionality to mine and analyze this large dataset. To present the result in the form of rich visualizations and to make better analysis, Business Intelligence tools have been used. The ability to mine and analyze big data gives organizations richer and deeper insights into business patterns and trends. The performance metrics of the proposed system is evaluated on the basis of execution time, data heterogeneity, scalability, flexibility and mining algorithm used.

[4] PR1169 DATA MINING TECHNIQUES FOR DIGITAL FORENSIC

By MASCARNES, SWEEDLE ANDREW

2014

005.74 MAS

Project Guide DR JOANNE GOMES

Abstract

Recently digital forensics has become a prominent activity in crime investigation since computers are increasingly used as tools to commit crimes. During forensic investigation the digital devices such as desktops, notebooks, smart phones etc. found at the crime scene are collected for further investigation. Investigators have to go through a humongous amount of data stored on these devices to gather evidence. This activity exceeds the expert's ability of analyzing and interpreting

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

the data. In this context data mining techniques such as clustering are used for automated data analysis. This research work focuses on a novel document clustering model that allows an investigator to semantically cluster the documents stored on a suspect's digital devices with the help of subject suggestions initially provided to him. Experiments were conducted on a dummy crime dataset to test the accuracy and the scalability of the proposed system. Experimental results proved that subject suggestion improved the accuracy and speeds up the process of searching the evidence. Without subject suggestion, the investigators may fail to give appropriate search query as they are completely heedless about the suspect's dataset and which may delay the process of investigation.

[5] PR1170 AUTHENTICATION USING MOBILE PHONE AS SECURITY
TOKEN

By KINI, MONALISA

2014

005.82 KIN

Project Guide MR.SHAMSUDDIN KHAN

Abstract

The inherent challenges of the security issue have become a top priority in every organization that makes use of information. Securing digital identities is getting more and more crucial. The usage of passwords for authentication is no longer sufficient because it faces most modern means of attacks and thus stronger authentication schemes are necessary. Strong authentication solutions using two identification factors often require an additional device, which could be inconvenient for the user and costly for the service providers. To avoid the usage of additional devices, the mobile phone is adopted as a security token. We use our mobile miss call as a password to the application. For a lot of security reasons one generally requires a very secure password, to implement the same the focus of this dissertation is on Authentication using mobile phone as security token and a mobile missed call is a unique one.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[6] PR1171 CONTENT-BASED MESSAGE FILTERING FROM ONLINE SOCIAL WALLS

By D'MELLO, SNEHAL WALTER

2014

005.82 DME

Project Guide DR A.K.SEN & MRS DAKSHATA PANCHAL

Abstract

Online Social Networks (OSNs) are very common in today's world. OSNs typically have an area where users post messages or comment on posts written on public / private areas commonly known as user walls. Users have the ability to view the public posts made by other users. However OSN users have no direct control on the content of messages posted on their walls, apart from blocking another user entirely from writing on his wall. Users may not be interested in viewing all the messages that are posted on their walls, and might not wish to block another user entirely from writing any message. In this work we give the users the ability to control the messages posted on their walls through customizable filtering rules applied to user walls. The messages are classified into different categories such as Sexual, Political, Religious and Vulgar using machine learning techniques like Radial Basis Probabilistic Neural Network and Radial Basis Function Network. Based on the output of the classification, the filtering rules decide the posting of messages on the user walls.

[7] PR1172 MOBILE DEVICE DATA SECURITY ON ANDROID

By RUMAO, VANESSA JOACHIM

2014

005.82 RUM

Project Guide DR A.K.SEN & MRS DAKSHATA PANCHAL

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Nowadays, mobile devices are an important part of our everyday lives since they enable us to access a large variety of ubiquitous services. Smartphones have become the most typical and popular mobile device in recent years. It combines the functionality of mobile phone and PDA. Besides, it provides many computer's functionality, such as processing, communication, data storage and etc. It also provides many computer's service, such as web browsers, portable media players, video calls, GPS, Wi-Fi and etc. In the same trend, the number and typologies of vulnerabilities exploiting these services and communication channels have increased as well. Therefore, smartphones may now represent an ideal target for malware writers. Therefore, Mobile devices, such as smartphones and tablets, typically need to support multiple security objectives: confidentiality, integrity, and availability. To achieve these objectives, mobile devices should be secured against a variety of threats. The proposed work provides the ability to lockout sensitive data, so that sensitive data should be protected from attackers. An access to a lockbox is based on time and location parameters and also it uses Graphical passwords which have been designed to try to make passwords more memorable and easier to use and, therefore, more secure. An important data is saved in an encrypted format. Using an encryption solution will protect the file from unauthorized viewing or modification as long as the encryption key is kept in the hands of authorized users. Therefore, encryption algorithm keys are stored on the server, so that unauthorized users can not access the keys and therefore important data. An encryption process uses a modified data encryption standard algorithm.

[8] PR1173 ENHANCED DATA PROCESSING USING POSITIVE NEGATIVE ASSOCIATION MINING AJAX DATA

By DOSHI, MUDRA RAJENDRABHA

2014

005.74 DOS

Project Guide MRS BIDISHA ROY

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Knowledge discovery is an important process in data mining wherein data can be analyzed from different perspectives and summarized for future use. One of the most widely used data mining processes is association rule mining. Association rule mining is a data mining process used widely in traditional databases to find the positive association rules. Association rules are created by analyzing data for frequent patterns and by using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the database. Confidence indicates the number of times the if/then statements have been found to be true. However, there are some other challenging rule mining topics like negative association rule mining. In this research, a rule mining approach has been proposed that provides an efficient solution using positive and negative association rule computation on Asynchronous JavaScript and XML (AJAX) data. AJAX is a new technique for creating better, faster, and more interactive web applications with the help of XML, HTML, CSS and Javascript. A Horizontal Tree Approach is proposed for efficient data processing. The performance of this approach is compared with Apriori and FP-Growth algorithms. By using AJAX, we get the search result in the form of semantic data. Whenever data search from the database is intended, the next possible word of the search will be made available.

[9] PR1174 WEB USAGE MINING FOR E-COMMERCE SITES

By LOPES, PRAJYOTI CAJETAN

2014

005.74 LOP

Project Guide MRS BIDISHA ROY

Abstract

Internet acts as a very rich source of information and this data is growing exponentially with time. Many E-commerce organizations rely on these websites to attract new customers and retain the existing ones. In order to achieve this goal web log files can be used that records customer's traversal patterns. This raw data can act as a rich source of information in discovering hidden patterns and extracting valuable knowledge. Our research work focuses on providing real time

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

recommendation to all the visitors of the web portal irrespective of being registered or unregistered. It makes use of traditional web usage mining phases for data preparation and preprocessing. Good quality input data is undeniably crucial to produce better results. Therefore various preprocessing tasks are used for data cleaning and session's generation. Three different approaches are proposed to provide item recommendation namely: product based, user based and action based rational recommendation approach respectively. In product based approach recommendations are provided to one time visitors and unregistered users. User based approach emphasizes on registered users. Action based rational recommendation techniques construct lexical patterns to provide dynamic recommendation as per changing user behavior. The proposed recommendation system tries to overcome the limitations of binary ratings and remarketing faced by recommendation systems in E-commerce sites. Effectiveness of the proposed system is evaluated by collecting real time E commerce data and testing it against the three approaches.

[10] PR1175 RECOMMENDATION ON EFFICIENT DATA EXTRACTION FOR WEB DATA

By PATIL, PRIYANKA MOHANDAS

2014

005.74 PAT

Project Guide DR A.K.SEN & MRS ANURADHA

Abstract

A web database is an organized listing of web pages, which can be queried or updated through World Wide Web (WWW). Web databases generate Query Result Pages (QRPs) in accordance to queries posted by users. Many applications necessitate the automatic extraction of data from these query result pages. The result from query result pages is very important for many web applications. This is because the QRP is a consolidation of related data which is in fields of databases. While trying to extract QRPs initially the webpage data which is in the HyperText Markup Language (HTML) format is converted to Extensible Markup Language (XML) format using web harvest tool. The data in XML format is semi-

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

structured or unstructured data. The alignment of this unstructured data into a structured format is done using cosine similarity approach. The aligned data is used for recommendation and text mining purposes.

[11] PR1176 EXPERT DISCOVERY AND INTERACTION USING MIXED SERVICES

By PATIL, SUSHANT PRASHANT

2014

006.76 PAT

Project Guide Ms,NAZNEEN ANSARI

Abstract

An expert system comprises connected experts that provide help and support services in a distributed association. To search for expertise, continuous interaction with expert systems is essential. In many of these systems, it is complicated for a user to determine whether an answer obtained is from a reliable and authenticated source. To overcome this problem, this study provides a new ranking algorithm i.e., “ExpertTop” algorithm that allows users to obtain accurate answers from reliable and authenticated sources. Moreover, this algorithm ranks the system experts through scores from the experience and feedback provided by co-experts and users, respectively. Naïve Bayesian algorithm is used for classifying a user’s query. Furthermore, ExpertTop algorithm will assist users in obtaining results from reliable and authenticated sources, which will provide them satisfactory experience.

[12] PR1177 PRIVACY PRESERVING DATA MINING FOR SOCIAL NETWORKS

By COLACO, BRINAL

2014

005.74 COL

Project Guide MR.SHAMSUDDIN KHAN

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

Advances in technology have made it possible to collect personal and professional data about individuals and the connections between them, such as their email correspondence and friendships on various internet social groups. In many cases the social network data describes relationships that are private and sharing the data for analysis can result in unacceptable disclosures. Online Social Networks are increasingly utilized by many users to publish details about themselves and to connect to their friends. Most of the information revealed inside these networks is not private. Yet it is possible to use learning algorithms on released data to predict private information from public information. This work focuses on the problem of private information leakage from the information present on the social networks. Five main attributes namely music interest, state of residence, groups, likes and friend list information are considered to infer religion, political affiliation and sexual orientation of the user. This inference is done using the Mamdani fuzzy inference system and the Naïve Bayesian classification. In order to prevent inference attacks on the social network data, the system suggests ways to hide or remove one or all of the five attributes mentioned. This process is termed as sanitization. Accuracy, precision and recall of the Mamdani Fuzzy inference system and Naïve Bayesian classification is computed. Comparing the Mamdani Fuzzy Inference System with the Naïve Bayesian technique on the same dataset, fuzzy inference systems gives better results.

[13] PR1178 PROTECTING DATA LEAKAGES IN CLOUD COMPUTING

By MANDHARE, SUPRIYA POPAT

2014

005.82 MAN

Project Guide DR A.K.SEN

Abstract

Popularity of cloud computing is increasing now-a-days. With the increasing use of cloud computing we can share information, resources, and data. Also we can store our personal as well as business information. Though it provides various advantages like pay-as-you-use, high flexibility, multi tenancy but with the new

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

computing and communication paradigms arises data security challenges. Existing data protection mechanisms failed in preventing data stealing attacks performed by an insider or hacker to the cloud provider. This research proposes a different approach for protecting data leakages in the cloud. In this work, various layers of security for data have been proposed which will provide highest security to data in the cloud. These layers will consist of user authentication with one time password, user access rights, encryption of users data and in the last layer hacker cannot access more information than the part which is stored in the cloud because system will automatically differentiates the file breaks and disjoint stores on different location

[14] PR1179 A METHODOICAL APPROACH FOR DETECTION AND 3-D RECONSTRUCTION OF BRAIN TUMOR IN MRI

By LOPES, SAYALI DOMINIC

2014

006.37 LOP

Project Guide DR DEEPAK JAYASWAL

Abstract

It always requires a skilled neurologist to detect a tumor in the MRI scans, which the neurologist does with the naked eye. This method is highly relying on the physician's prowess and ability to discern between the normal and abnormal brain scans. A machine being highly complex can be exploited to do this work with great amount of repeatability and accuracy without the factor of fatigue. Also all these years it has been observed that Doctors have had only the 2D Image Data to visualize the tumors in the MRI images which never gave the actual feel of how the tumor would exactly look like. Since the doctors were deprived from the exact visualization of the tumor the amount of the tumor to be removed by operation was not known, which caused a lot of deformation in the faces and structure of the patients face or skull. This research presents a method for automatic tumor detection of tumors with an added feature of reconstructing its 3D image, one of the most attractive avenues in digital image processing techniques, especially due to its application in biomedical imaging. The diversity and complexity of tumor

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

cells makes it very challenging to visualize tumors present in magnetic resonance image (MRI) data. The research involves implementation of various steps for detecting and extracting the tumor from the 2D slices of MRI brain images by Seeded region growing technique along with automatic seed selection and designing software for reconstructing 3D image from a set of 2D tumor images. The seeded region growing method is a very attractive method for semantic image segmentation which involves high level knowledge of image components during the seed selection procedure. Extensive use of custom made user interface that provides for ease of user interaction and visualization of reconstructed data. The volume of the tumor is also estimated based on the computation of these images. Doctors and Radiologists can now prepare and image thousands of samples and save on time per day using this automation

[15] PR1184 DIGITAL RIGHTS MANAGEMENT MECHANISM WITHOUT THIRD PARTY

By D'MELLO, SHERYL RAPHAEL

2014

005.82 DME

Project Guide MRS BIDISHA ROY

Abstract

The advance in computing and easy availability of technology has led to an increase in the consumption of digital content. It has also led to illegal replication and distribution of digital contents. This poses a huge threat to organizations dealing with highly confidential digital contents and could drastically affect its financial standing. Digital Rights Management (DRM) technologies aim to provide protection and secure distribution of the digital contents. This research proposes a different DRM mechanism for protecting and securing the distribution of digital content within an organization. It aims to benefit the organizations by providing security, privacy, accountability as well revocation of malicious users. It eliminates the need for third parties by using the system details to provide controlled access and prevent illegal redistribution of the digital content. Various layers of security

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

have been proposed at each level that enhance the security of digital content and simultaneously achieve privacy, accountability and revocation of a malicious user without the use of third parties. These layers consist of user authentication, account activation and change system feature using one time password, digital content access rights using system verification process, encryption and decryption technique to secure digital content, hashing technique to provide accountability as well as to maintain the privacy of the user. The accountability feature of the proposed mechanism enables to track the illegal attempts made to access the digital content. The revocation feature of the proposed mechanism further disables the access to the digital content.

[16] PR1189 LOAD REBALANCING ON CLOUD

By MEHER, ANAGHA MANOHAR

2014

006.76 MEH

Project Guide MRS BIDISHA ROY

Abstract

Cloud computing is distributed computing over a network and it means the ability to run a program on many connected computers at the same time. Large scale distributed systems such as cloud computing applications are becoming very common. These applications come with increasing challenges on how to transfer and where to store and compute data. Load balancing is one of the challenging tasks. Load balancing is the process of reassigning the total loads to the individual nodes of the collective system to make the best response time and also good utilization of the resources and to remove the situation where some nodes are overloaded and some others are under loaded. To decrease the total number of heavy nodes (servers) in the system by moving load from heavy nodes (servers) to light nodes (servers) is the main aim of balancing the load. Our objective is to allocate the processes as uniformly as possible among the nodes such that no node manages an excessive number of loads. In this paper we present K-means algorithm,

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Min-Min and Max-Min algorithm for load balancing on cloud.

[17] PR1190 PREDICTION AND ANALYSIS OF DISEASE THROUGH
DATA MINING TECHNIQUES

By PAVATE,ARUNA ANIMISH

2015

005.74 PAV

Project Guide Ms,NAZNEEN ANSARI

Abstract

Diabetes has become the fourth leading cause of death in developed countries. By the endurance and hasty spread of diabetes, with increased number of ill health, complications in the disease all over the world, several methodologies have been developed to predict and prevent this chronic disease. An early diagnosis of disease helps patients and medical experts to reduce the problem, risk and cost of medications. This work presents the design and development of an efficient system to predict diabetes and its further complications with risk level. In this system, three methods including genetic algorithm, nearest neighbor, and fuzzy rule-based system have been used in order to provide an accurate prediction system to prepare for the presence of diabetes. In this system, 530 individuals were involved and their records were collected. The best subsets of features generated by the implemented algorithm include the most common risk factors such as age, family history, BMI, weight, smoking habit, alcohol habit and also factors related to the presence of other diabetes complications which were used for prediction of disease. The proposed system was prejudiced and the results showed to be more suitable by selecting the best subset of features selected using variations of genetic algorithms. The system shows that accuracy of prediction of disease improves when the number of selected features and number of patients records increases. The succeeded results produced 95.83% sensitivity, 95.50% accuracy and 86.95% specificity on impenetrable data which support the effectiveness of the system to predict the disease.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[18] PR1191 PREDICTION OF FOREX RATES USING HYBRID SYSTEM

By D LIMA,NATHAN

2015

DLI

Project Guide MR.SHAMSUDDIN KHAN

Abstract

Currency exchange is the trading of one currency against another. Forecasting the foreign exchange rate is an uphill task. FOREX rates are influenced by many correlated economic, political and psychological factors. Like many economic time series, FOREX has its own trend, cycle and irregularity. Many methods have been used to predict the future FOREX rates. Some of the common methods include statistical analysis, time series analysis, etc. Some of the modern methods used are fuzzy systems, neural networks, hybrid systems, etc. These methods suffer from the problem of accurately predicting the exchange rates with low accuracy and precision. The objective is to predict future exchange rates for the next 15 days with higher accuracy and precision. By using two different methods, i.e. a Neural network and a Hybrid system, a more accurate and robust method can be developed. An Artificial Neural Network (ANN) is used to predict the rise and fall of the FOREX market while an ANFIS system is used to predict the future rate. The results are evaluated in terms of Mean Square Error (MSE) and Mean Absolute Error (MAE).

[19] PR1192 OBESTDD:ONTOLOGY BASED EXPERT SYSTEM FOR THYROID DIAGNOSIS

By RAWTE,VIPULA DATTATEREY

2015

006.32 RAW

Project Guide MRS BIDISHA ROY

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

The thyroid gland is one of the most important organs in the body since it is responsible for controlling metabolism. Thyroid diseases are mainly classified into two type viz., hyperthyroidism and hypothyroidism and these can be difficult to diagnose because symptoms are often easily confused with other conditions. Expert system in medical field are very helpful since it can process huge amount of data which otherwise is cumbersome. Such traditional expert systems use machine learning techniques which require training of data and incurs time. Ontology represents explicit specification of knowledge in a specific domain of interest in the form of concepts and relationships among them. It has an advantage of being both computer and human readable. The adoption of ontology in healthcare has facilitated domain experts and non-experts to perform knowledge representation tasks with great ease.

An ontology describing the domain of thyroid diseases and its symptoms is presented. This study proposes at diagnosing thyroid diseases with an ontology based expert system called as OBESTDD (Ontology Based Expert System for Thyroid Disease Diagnosis). It uses ontology to model the contents of domain knowledge and rules to infer the related diagnosis.

[20] PR1193 TAXI-OUT TIME PREDICTION

By GEORGE,ELIZABETH

2015

519.3 GEO

Project Guide MR.SHAMSUDDIN KHAN

Abstract

Flights undergo a large percentage of delay between scheduled departure of an aircraft and actual takeoff. This not only leads to passenger resentment but also, results in emission of harmful pollutants that adversely affect the environment. The major causes of delays in the on-time performance of flights are due to air traffic congestion, congestion at the departure terminal, gate push-back delay (taxi-out time delay), other reactionary delays, etc. Delays in taxi-out time have a noteworthy impact on the airline's economy and public health. Also, since the airport operations are dynamic in nature, prediction of taxi-out times for all flights

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

can be challenging. The aim is to show that this method provides promising and potential features to tackle the airport departure problems. This research tests three approaches: Linear Regression, ANFIS & Q-learning, which fall under the realm of prediction to accurately predict the taxi-out times. Historic data of an airport is analyzed and utilized for validation. A novel Q-learning algorithm is proposed to predict the accurate taxi-out times at a specific airport. Operational data is analyzed using the Markov Decision Process (MDP) after which a reinforcement learning methodology for Q-learning is formalized for the estimation of taxi-out time of flights. The predicted taxi-out time result is then compared with the actual taxi-out time of the next set of data. This is done to reduce the taxi-out time error and thus by correctly predicting the taxi-out time harmful emissions and other reactionary delays are reduced. The mean square error obtained from Q-learning, Linear Regression and ANFIS are calculated to identify which of these methods have the least prediction error for taxi-out time.

[21] PR1194 AD-HOC VEHICLE ROUTING PROBLEM

By CARWALO, TEJAL FRANCIS

2015

388.3 CAR

Project Guide MRS VANDANA PATIL

Abstract

In a vehicle routing problem (VRP), there are one or more vehicles that travel around a network, which depart from and return to a start node. This study aims to minimize the total route length while considering nodal demand (capacity constraint) of the vehicle as well as considering the dynamic addition and deletion of nodes. The ant colony optimization (ACO) algorithm is an effective algorithm for solving the problem of optimizing a shortest route. ACO is a meta-heuristic algorithm that has been widely used for different combinatorial optimization problems and inspired by the foraging behavior of real ant colonies. For a large area network, the use of a clockwise partition clustering algorithm can significantly improve the solution efficiency. This study proposes the partition clustering to divide the whole area into the different clusters while considering nodal demand of

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

the vehicle and then it uses ACO algorithm on each cluster to find the shortest path so that travelled distance can be minimized.

[22] PR1195 MULTIPLE TEXT DOCUMENTS SUMMARIZATION SYSTEM

By DAVE, HARSHA ASHOK

2015

003.35133 DAV

Project Guide MRS SHREE JASWAL

Abstract

Text Summarization plays an important role in the area of text mining and natural language processing. As the information resources are increasing tremendously, readers are overloaded with loads of information. Finding out the relevant data and manually summarizing it in a short time is a much more difficult, challenging and tedious task for a human being. Text Summarization aims to compress the source text into a shorter and concise form with preserving its information content and overall meaning. Summarization can be classified into two main categories i.e. extractive summarization and abstractive summarization. This paper presents a novel approach to generate abstractive summary from extractive summary using WordNet ontology. An experimental result shows the generated summary in well-compressed, grammatically correct and human readable format.

[23] PR1196 CONTEXTUAL QUERY EXPANSION FOR ACQUIRING WEB DOCUMENTS

By UPADHYAY, PRIYANKA ASHOK

2015

005.74 UPA

Project Guide MRS BIDISHA ROY

Abstract :

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

In today's world, Contextual Information retrieval over the internet has become a major issue. Query expansion is an information retrieval technique in which new query terms are added to the original query terms to improve search performance. Contextual query expansion is a major issue in today's era. In this research work, contextualization is achieved by performing document extraction and terms extraction activities to the particular domain information source. User query is expanded using document extraction and terms extraction activities. Document extraction is achieved by BM25 retrieval function. It ranks a set of documents based on query terms appearing in each document. Now the second function is terms extraction process, in this, terms in the top returned documents are weighted using sub linear terms frequency scaling formula which is used to weight the terms in the expanded query derived from original query which will cope with the term mismatch problem in specific domain. This research work focuses on a proposal to make web searches adaptive to the context of the user's query, thus improving query results. The proposed approach makes the contextual information faster. In addition the results of the query engine with and without the contextual information showed improvements in the precision and decreases search length of the web results.

[24] PR1328 DETECTION OF WEB DESIGN PATTERNS USING REVERSE ENGINEERING

By THANKAPPAN, JERIN

2015

006.76 THA

Project Guide MRS VANDANA PATIL

Abstract

Good use of design patterns increases code maintainability and provides good understanding to other developers. Design Patterns are the most valuable and approachable solution for any software design issues. It is one of the crucial tasks in software comprehension. Pattern identification delivers important information to the designers. The proposed work here

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

gives a detailed analysis of this solution and also helps in fulfilling the detection of design patterns in java. Since design is not tangible, it is hard to detect. The proposed application is functioned for all kinds of java based software where any person interested in understanding the code can use. The focus here is to detect java design patterns and to give a detailed understanding of the system by producing a proper GUI which displays classes and methods in each class and also by showing the inheritance or implementation by general convention. Functionally, this tool gives an overall idea of the design pattern and its detection. Our approach for this tool is for some behavioral patterns which include Observer, Composite, Mediator, State, Template, Command, Decorator, Factory, Proxy, and Singleton and if the Classifier does not recognize the pattern then it only displays the tree diagram. This is achieved by using a new method of detection using similarity scoring for graphs. It helps in reducing the efforts in the learning curve when dealing with a particular piece of code. By reverse engineering these applications we can make it possible for other (secondary) developers to understand the code better which will reduce human efforts of maintaining any documents for the Web Application.

[25] PR1329 RELEVANT SEARCH RETRIEVAL ON WEB

By RODRIGUES,LISSA PETER

2015

025.04 ROD

Project Guide MRS SHREE JASWAL

Abstract

Nowadays, the web has become widespread in terms of availability of contents related to every field. Also a large repository of web contents is turned up as a most challenging tool for searching and retrieving information. For scientists and researchers, resource or content searching has been very important. Today's market is full of variant search tools over the web having discrepancies in terms of working and the end search results. Search tools normally return a huge number of

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

relevant web pages, for a given query. This study presents a hybrid approach of Enhanced-Ratio Rank and Page level keyword algorithms in order to get most relevant pages at top of search results.

[26] PR1331 SECURE DATA TRANSMISSION USING VIDEO

By LEMOS, NIKITA NORMAN

2015

005.82 LEM

Project Guide MRS BIDISHA ROY

Abstract

In today's world transmitting data in a safe and secure fashion is difficult, especially when highly sensitive data is involved. The data should be ideally robust to security threats. The dissertation work aims at proposing a methodology which employs dual level of security using cryptography and steganography to hide the secret text data using video as a cover. A blend of existing and novel techniques is used for hiding the data. Cryptography aims at concealing the secret data and steganography hides the existence of the data. Since a video is addressed as a collection of frames, a frame selection logic is incorporated which inserts data into the frames in a random fashion so that the data cannot be retrieved easily by the attacker. Considering the goal of cryptography and steganography the analysis of the methodology can be done on the basis of visual perceptibility, error ratios and histogram comparison of video frames before and after hiding data. Use of two techniques with distinct goals to hide the data increases the robustness as well as simple frame selection logic prevents the attacker to obtain the data easily but at the same time provides ease of use

[27] PR1336 HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES

By D BRITTO, ROVINA REGAN

2015

610.285 DBR

Project Guide MRS ANURADHA

ST. FRANCIS INSTITUTE OF TECHNOLOGY (ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

A common term Heart Disease is nothing but a cardiovascular disease or a Coronary heart disease which reduces the efficiency and proper functioning of heart by blocking veins, arteries or blood vessels around it. Coronary heart disease causes disability such as damage to the brain resulting in death. About 25% of deaths (age group 25-69 years) occur because of heart diseases. The reasons for cardiovascular disease are smoking, physical inactivity, an unhealthy diet and harmful use of alcohol which can result in chest pain, stroke and heart attack. Since the factors and symptoms that are responsible for heart problems are known, it is possible to predict any heart problem based on statistical analysis done on the basis of medical records. However Data mining techniques provide an automatic way of analyzing the data using standard classification methods. There are many classifiers available in data mining for predicting heart disease. This thesis emphasizes on finding the appropriate classifier that will give better accuracy. The three data mining techniques used here are Naïve Bayes, Support Vector machine and Logistic Regression.

[28] PR1339 SCIENTIFIC DOCUMENT CLASSIFICATION

By D CUNHA,ARLINA ANTHONY

2015

005.133 DCU

Project Guide DR A.K.SEN & MRS VINCY JOSEPH

Abstract

Classification is the alignment of data or items in predefined labelled groups based on resemblances. Exponential progression of scientific documents leads to uncontrollable physical classification. Feature extraction is the crucial condition of automatic document classification. TF-IDF (term frequency-inverse document frequency) is frequently used to represent the text feature weight due to its ease of computation. This report proposes a new yet simple feature weighting scheme by modifying TF-IDF formula. In the proposed method, feature weight is calculated according to the distribution of terms inside a document and in different

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

positioning levels; as terms appearing in higher structural level of the document are more significant than terms appearing in lower structural levels. The experimental results show that the modified method improves the accuracy, execution time and other parameters.

[29] PR1340 FATIGUE DETECTION SYSTEM USING IMAGE PROCESSING ON VIDEO SEQUENCES

By KHAN, NEEHA

573.85 KHA

Project Guide MRS DAKSHATA PANCHAL

Abstract

Driver drowsiness problem is one of the important reasons that create road accidents. The system detects fatigue based on parameters such as eye state detection yawning state head rolls and head tilts. In such a case when drowsiness is detected, an alarm signal is issued to alert the driver. Then yawning detection is also performed to determine the drowsy state using various operators. Developing intelligent systems to prevent car accidents can be very effective in minimizing accident death toll as one of the factors which play an important role in accidents is the human errors including driving fatigue. To address the problem a vision-based driver fatigue detection system is developed which considers video samples which represent a real time running scenario. Eye tracking is one of the key features for driver assistance systems since human eyes contain much information about the driver's condition such as gaze, attention level, and fatigue level. The algorithms used are Adaboost Algorithm and various edge detection Algorithms such as Canny Algorithm and Sobel Algorithm. Face and eyes of the drivers are first localized and then marked in every frame obtained from the video source. Along with the eyes the mouth and the head rolls also play a major role in detecting fatigue as considering only the eye region won't make the system stronger and more efficient. This system can minimize the number of accidents caused by driver fatigue. Driver fatigue is a result of sleep deprivation or sleep disorders is an important factor in the increasing number of accidents on today's

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

roads. The purpose of this project was to advance a system to detect fatigue symptoms in drivers and produce timely warnings that could prevent accidents. Thus a system that can detect oncoming driver fatigue and issue timely warnings could help prevent many accidents, and consequently save money and reduce personal suffering.

[30] PR1342 BIOMETRIC IDENTIFICATION USING IRIS

By PATIL,ARTI PRABHAKAR

2015

006.4 PAT

Project Guide DR DEEPAK JAYASWAL

Abstract

The recent advantages of the information technology and the increase of the security requests have entailed a fast development of biometric technique based personal identification smart systems. Iris recognition and authentication is a very recent technique in the area of personal identification. Iris recognition system is a reliable and an accurate biometric system. Localization of the iris borders in an eye image can be considered as a vital step in the iris recognition process. There exist many algorithms to segment the iris. One of the segmentation methods that is used in many commercial iris biometric systems is an algorithm known as a Daugman's algorithm. The aim of this thesis is to implement this algorithm using the MATLAB programming environment. The implemented algorithm was tested on the eye images of different quality, such as countering the effect of light reflection on images another quality is the images with partly covered iris or low contrast and noisy images. If a person is suffering from light reflections on eye images affects the iris detection so that original information of this part of the image has been lost and also there is an observable difference in intensity values between 'light affected region' and its surrounding darker region. In order to avoid light reflection affecting the iris detection, a morphological operator is used. The test results demonstrated that the Daugman's algorithm detects the iris borders in the high quality images with high accuracy. The performance of the algorithm on the

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

lower quality images has been improved by additional preprocessing of these images.

[31] PR1345 MINING COMPUTER GAME DATA FOR BENEFITING GAMING INDUSTRY

By ANDRAT, HYACINTA

2016

794.811526 AND

Project Guide Ms NAZNEEN ANSARI

Abstract

Playing computer games for many years has led to a large volume of gaming data that consist of gamers' likings and their playing behavior. Such data can be used by game creators to extract knowledge for enhancing games. Mining computer game data is a new data mining approach that can help in developing games as per a gamer's requirements and his/her area of interest. Since the gaming industry has been contributing to the countries' revenue on a large scale, improvement in this industry becomes vital. This study aims to apply data mining techniques such as association, classification, and clustering for improving game design, game marketing, and game stickiness monitoring, respectively, to enrich game quality.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[32] PR1346 AN APPROACH FOR DIAGNOSIS OF ANXIETY
DISORDER

By D MONTE, SILVIYA

2015

616.8522 DMO

Project Guide MRS DAKSHATA PANCHAL

Abstract

The mental Disorder is one of the top five reasons for the cause of death in world and Suicide is the third leading cause of death among young adults worldwide. Due to the unawareness and social Embarrassment, this disease is normally undiagnosed. Study says that one of every four Indians affected by anxiety disorder, out of which 10% depressed. The reasons affecting the mental health of the Indian patients need to be identified through which better cognitive behavioral therapy can be given by the psychiatrist. Data mining is a way through which Long term monitoring of the patient can be carried out by using stress monitoring tests and actual reasons affecting the mental health of the Indian patient can be studied via digging through and analyzing enormous sets of data and then extracting the meaning of the data such as their lifestyle, habits etc. It will also provide healthcare professionals an additional source of knowledge for making decisions. The aim of this research is to propose a new approach using data mining techniques to predict the stress level of a patient using logistic model trees and to find out different factors affecting the mental health of a Indian patient in an efficiently and an economically faster manner..

[33] PR1348 REALIZING STUDENTS' UNDERSTANDING THROUGH RULE
BASED REASONING

By DOSHI, VRAJESH PANKAJ

2016

006.74 DOS

Project Guide MRS BIDISHA ROY

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

Lectures are one of the mediums used in Teaching-Learning strategy. However, the amount of knowledge gained by the student is not always equal to the amount of knowledge shared in the lecture. This can be due to several reasons like long lecture hours, lack of concentration of students and many others. If the concepts which are not clearly understood are prerequisite to understand the next/other concept/topic then there are high chances that students won't understand even that. In such scenarios, to ensure that all students have understood each and every concept taught, there is a need for identifying which concepts the students have not understood properly. So to fulfill the requirement, an online exam evaluation system is proposed that makes use of a customized online exam system and generates custom reports using Machine Learning and Rule Based Reasoning. These custom reports make the system useful not only for assessing the proficiency of students' but also help the teacher in realizing their true state of understanding for the knowledge shared in class. Keywords- Online Examination System, Evaluation, Assessment, Student, Lecture, Understanding, teaching, learning, Multiple Choice Question (MCQ), PARAM.

[34] PR1349 CO-CLUSTERING WITH SIDE INFORMATION FOR TEXT MINING

By THOMAS, RAMYA ELIZABETH

2016

005.74068 THO

Project Guide MR SHAMSUDDIN KHAN

Abstract

Text mining applications can obtain a huge amount of information from documents in the form of text. This text information can prove as an input for Text Clustering. Along with text in these documents, it also includes various kinds of other information known as Side Information or Metadata. Examples of this side information include links to other web pages, title of the document, metadata about

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

the document, author name or date of Publication which are present in the text document. Such metadata may possess a lot of information for the clustering purposes. But this Side information may be sometimes noisy. Using such Side Information for producing clusters without filtering it, can result in a bad quality of Clusters. This study proposes an efficient Feature Selection method to perform the mining process to select that Side Information which is useful for Clustering so as to maximize the advantages from using it and makes use of the process of Co-Clustering or Two-mode clustering which is a data mining technique that allows producing groups by Clustering both Text and Side Information.

[35] PR1350 ANALYSING TRUSTWORTHINESS IN B2C E-COMMERCE WEBSITES

By BRAHMBHATT, MANTHAN YOGESH

2015

658.84 BRA

Project Guide MRS BIDISHA ROY & MRS PRADNYA RANE

Abstract

Information system professionals must pay attention towards retaining and alluring potential customers for e-commerce vendors. With the increase in the number of internet users, the customers for online shopping have also increased considerably; so for any e-commerce vendor it is really important to retain the customers. Retaining and alluring these potential customers can be done only when customers trust the e-commerce vendors. Trust lies at the acme in this highly competitive trend of e-commerce. Trust is a very important factor in any social relationship and especially in commerce transactions. So analyzing those trust factors is also very important. So in this work, we analyze trust using different approaches like ANN approach, ANFIS approach and Text Analysis approach. The objective is to compare all these approaches and get the results about which approach is more suitable for analyzing trustworthiness in B2C e-commerce websites. These comparisons are done on the basis of Mean Square Error (MSE) and time. Apart from these applications based on all these approaches are also discussed.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[36] PR1354 FUZZY COGNITIVE MAPS FOR OBESITY
PREDICTION

By NADAR,NANCY

2016

511.313 NAD

Project Guide MRS ANURADHA

Abstract

The World Health Organization has described Obesity as one of today's most neglected public health problems, affecting every region of the globe. Obesity has become the concern in many countries. There are many factors that contribute to the occurrence of Obesity. The government indulgent regulation on food restriction provides easy accessibility to unhealthy, processed food. Also India's growing career oriented lifestyle gives rise to irregular biological patterns especially in the younger generation who prefer indoor games rather than playing outside. The behavioral and socio-psychological factors also contribute to obesity in a person. Other factors that influence are sleep, stress, Ethnicity, among many other multiple such factors which give rise to obesity in today's era. Fuzzy Cognitive Maps constitute an attractive knowledge-based methodology, combining the robust properties of fuzzy logic and neural networks. Medical Systems are complex systems taking into consideration many factors that may be complementary, contradictory, and competitive. These factors influence each other and determine the overall diagnosis with a different degree. These systems also involve inexact, uncertain, imprecise and ambiguous information. The Information available may be vague, missing or not available. FCM can handle the interactions between the concepts which are vague and not always concrete. FCMs can even handle incomplete or conflicting information. This factor is especially important in medicine where experts should take many factors under consideration before they can reach a decision.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Hence the proposed work uses Fuzzy Cognitive Map to develop a mathematical model that provides estimates of probability for obesity in a person so that timely preventive action can be taken.

[37] PR1356 WEB USAGE MINING USING HIERARCHICAL CLUSTERING

By SHAH, FORAM

2015

005.74 SHA

Project Guide DR JOANNE GOMES

Abstract

Web usage mining is a branch of web mining in which by clustering the datasets, frequently accessed patterns can be obtained for betterment of social portals, websites. The divisive analysis is one of the types of hierarchical methods of data clustering that is used to separate each dataset from the clustered data depending on previous small clusters. Apriori & k- Apriori are commonly used hierarchical algorithms for web usage mining but they are less efficient for mining dynamic item sets such as twitter dataset. Recently D-Apriori algorithm has been proposed in the literature for mining dynamic dataset to find the frequently accessed web pages from web log databases. The disadvantage of D-Apriori is that it takes more execution time compared to k-Apriori. This research work proposes a new hybrid k-D-Apriori algorithm that improves frequent pattern generation, works efficiently with dynamic datasets and gives improved association rule generation as compared to D Apriori algorithm.

[38] PR1359 USER NAVIGATION PATTERN PREDICTION

By JANISA, JOHN COLACO

2016

005.74 JAN

Project Guide MRS JAYASHRI MITTAL

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

User navigation pattern discovery is a technique of learning the user's browsing pattern while he is navigating on the web. It is mainly concerned with web usage mining. As the web log files contain an entire record of user's browsing behavior, such log files can be analyzed to identify the user's navigation pattern. The extracted navigation patterns are further used to predict the user's future request. In this study hybrid approach is proposed based on the K- harmonic means algorithm for clustering and the frequent pattern growth algorithm for identifying frequent patterns from it. Such frequent patterns are further used to predict the URL's likely to be visited by the user, based on a recent URL the user has visited. The applicability of this approach will be in several areas like recommendation, prediction, web caching policies, personalization etc.

[39] PR1496 PRIVACY PRESERVING IN DATA MINING

By FERNANDES,MELISSA

2016

005.74068 FER

Project Guide DR JOANNE GOMES

Abstract

In many organizations huge amounts of data is generated. Organizations use this data for their own benefit. Data mining extracts useful knowledge from huge data. Association rule mining is a powerful technique to find hidden patterns in large databases. The limitation of mining association rules is that some sensitive patterns are revealed from sensitive rules. It is necessary to hide sensitive rules due to privacy concerns. The technique used to hide sensitive rules in order to get sanitized data is called Association rule hiding. The paper aims at proposing a methodology which includes a blend of Equivalent Class Transformation (ECLAT) algorithm used to find frequent itemsets and heuristic approach which is one of the methods of Association rule hiding. The proposed hybrid approach is evaluated on the basis of

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

execution time, dissimilarity, lost rules, hiding failure and number of transactions modified.

[40] PR1497 IMPROVED WEB SERVICE RECOMMENDATION VIA EXPLOITING LOCATION AND QOS

By GONSALVES,BLESSINA BAPTISTA

2016

006.76 GON

Project Guide MRS VANDANA PATIL

Abstract

Web services describe a standardized way of integrating web-based applications that help in machine-to-machine interaction over the network. There are many publicly available web services and the number keeps on increasing. However this ever increasing pool makes it difficult for optimal service selection. So, appropriate selection of web service suiting the requirements of the user is a non-trivial task. Our research proposes a technique which helps in optimal services selection with optimal Quality-of-Service (QoS) performance. The technique designs a recommender system based on Collaborative Filtering (CF) algorithm which employs location and QoS values to cluster users and services. The main objective of the proposed technique is to address the issue of data sparsity and scalability. The proposed approach uses k-Nearest Neighbor (k-NN) algorithm with Support Vector Machine (SVM) in CF algorithm framework. SVM a state-of-the-art classification algorithm used to address the issue of sparse data and k-NN used with CF algorithm for similarity mapping of user and services.

[41] PR1500 ROBUST STEGANOGRAPHY TECHNIQUES FOR HIDING DATA IN IMAGES

By JOSHI,SHWETA

2016

005.8 JOS

Project Guide MR SHAMSUDDIN S KHAN

ST. FRANCIS INSTITUTE OF TECHNOLOGY (ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

The security of information is crucial in the digital world as the information may involve confidential data like credit/debit card number, ATM Pin, etc. Steganography is one of the techniques for securing confidential information by hiding it in some seemingly innocent cover medium (text, image, audio, and video). Steganography provides invisible communication and hides the existence of information. The steganography technique that uses digital images as a cover medium is called „Image Steganography“. The proposed work focuses on „before embedding technique“ of hiding in image steganography. It tries to find suitable places in the cover image to embed the secret image securely. In the proposed method, Genetic Algorithm is applied to identify appropriate places in the cover image where embedding of a secret image will cause minimum distortion. After these places are obtained, embedding is performed using transform domain techniques (Discrete Cosine Transform, Discrete Sine Transform and Kekre’s Transform) and their corresponding Wavelet Transforms. Hybrid Wavelets like Discrete Sine Transform-Discrete Cosine Transform (DST-DCT) and Kekre’s Transform-Discrete Cosine Transform (KT-DCT) are also applied. Then the secret image is embedded in the lower energy transformed blocks of the selected cover image regions to further improve the stego image quality. All the proposed algorithms are experimented with thirty cover images of the same size and six secret images of variable sizes. All algorithms are evaluated and compared using the same set of performance evaluation parameters like Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Correlation and Structural Similarity Index Measurement (SSIM). The proposed system gives new direction in image steganography by providing the role of genetic algorithms and different transform domain techniques in selecting the suitable blocks of cover image for hiding secret images that leads to improved security along with increased embedding capacity while reducing the computational complexity.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[42] PR1501 RECAPITULIZATION OF TWEETS USING GRAPH-BASED CLUSTERING

By LOBO, VIVIAN BRIAN

2017

005.741 LOB

Project Guide Ms NAZNEEN ANSARI

Abstract

Twitter—a well-liked online social networking site—facilitates millions of users on a daily basis to dispatch and orate quick 140-character notes named tweets. Nowadays, Twitter is cited as the fastest and popular intermediate method of communication and is used to follow the latest events. Tweets pertaining to a specific event can be effortlessly found using keyword matching, but there are numerous tweets that are likely to contain information that is semantically identical. Moreover, there exist many systems for recapitulating tweets related to a particular event, but they have numerous limitations and are unable to provide accurate results. This research work aims to overcome the limitations of existing systems by developing a system for recapitulating tweets using graph-based clustering. We evaluate our developed system via user ratings and show that the developed system outperforms several state-of-the-art recapitalization systems.

[43] PR1502 FRAUD DETECTION IN E COMMERCE TRANSACTIONS

By SHAJI, JISHA

2016

658.84 SHA

Project Guide MRS DAKSHATA PANCHAL

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Due to the advancement in technology the number of fraudulent activities is increasing dramatically in every field. Fraud is a major problem in electronic payment systems. There has been a tremendous rise in electronic commerce in the recent years particularly due to online retailers like Amazon, Flipchart, eBay, AliExpress.com to name a few. Fraud begins to rise as new technologies and new weaknesses are found. As a result the number of fraud cases registered has also increased thus resulting in tremendous losses each year. Since the transactions associated with e-commerce are large in number, the dataset associated with them is also large; therefore it requires fast and efficient algorithms to identify fraudulent transactions. Reducing fraud is a complex process which includes the knowledge from many scientific areas and demands a multi-disciplinary approach. Many methods have been used to detect fraud in electronic transactions, especially credit card based transactions. Most of them are rule based systems or systems that require retraining as and when newer patterns of fraud occur. Detecting fraud as it is happening or within a short time span is not easy and requires advanced techniques. As the demand has arisen for self-learning predictive systems, the main objective is to detect fraudulent transactions by using Adaptive Neuro Fuzzy Inference System. It is a hybrid of neural networks along with fuzzy inference, wherein we can specify our own rule sets so that the system can adapt to newer instances of fraud.

[44] PR1503 NEGOTIATION AGENT OVER E COMMERCE SYSTEMS

By RODRIGUES,SALY SEBASTIAN

2016

658.84 ROD

Project Guide : MRS BIDISHA ROY

Abstract : Now-a-days, negotiation over e-commerce platforms is gaining popularity in the research field. Electronic markets are evolving in such a way that they could provide various products and services automatically over the Internet. Even though ample of services are made available over e-commerce systems, but the basic need of providing products and services that actually benefits the customer remains as it is. Dealers make a lot of deals, offers, provide coupons over the sites such as ebay.com, but there is less or no input

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

taken from the customers regarding their terms and conditions for a product while making the purchase. So bargaining over a product/service becomes a potential issue that would mutually reach an agreement which is beneficial for both the seller and buyer. There are a number of models proposed for negotiation. Here, proposed research work uses Adaptive Neuro Fuzzy Inference System (ANFIS) for classifying the user needs to negotiate over a product, and predicting a counter offer as a seller, to the offer made by the buyer. It then compares the results with Pareto Optimal (PO) solution gained by using Maximum Greedy Trade-off (MGT) Algorithm and a Multi- strategy Selection Algorithm to negotiate. ANFIS includes benefits of both Artificial Neural Network (ANN) and Fuzzy Logic (FL) Systems. And based on the comprehensive set of features and fuzzy rules the negotiation is carried out between two parties. Experimental results illustrate promising results. And a comparative analysis is performed with MGT and Multi-strategy Selection Algorithms to show the superior nature of ANFIS based negotiation model.

[45] PR1505 OPTIMIZATION OF SUPPLY CHAIN MANAGEMENT PROCESS

By TAWDE, PRACHI VASUEV

2016

658.5 TAW

Project Guide MRS SHREE JASWAL

Abstract

Demand Forecasting is estimation of demand depending upon the sales data or the previous demand data. Forecasting demand accurately is a herculean task, Demand Forecast depends upon factors such as the seasonal trend, price elasticity of product. Supply Chain Management(SCM) process includes tasks like procuring goods, inventory, and manufacture, planning logistics, facility location, shipping and dissemination. All these functions are affected in the short run by product demand and in the long run by products and processes and fluctuating

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

markets. Forecast of product determines quantity of product to be made and how much material to purchase from suppliers to meet forecasted customers needs. The objective is to predict the demand on a daily basis taking previous two days data to predict the third day. By using Support Vector Regression (SVR) for predicting the demand and optimizing its parameters using Artificial Bee Colony(ABC).

[46] PR1508 CREATING ROUTING PLAN FOR KEYWORD QUERY

By NUNES,PRANALI DENIS

2016

004 NUN

Project Guide MRS ANURADHA

Abstract

Keyword search is an inherent example for searching linked data sources on the web. The proposed work tries to route keywords only to relevant sources. This would also reduce the high cost of processing keyword search queries over all sources. This proposed work also tries to compute top-k routing places based on their potentials to contain results for a given keyword query. The number of potential results may increase exponentially with the number of sources and links between them. Most of the results may be not necessary especially when they are not relevant to the user. The proposed work uses keyword search to solve the problem of keyword search over large number of linked and structured data source by using keyword expansion.

[47] PR1515 DYNAMIC QUERY FORMS USING RANKING MODELS FOR DATABASE QUERIES

By KOSHTI,SPHURTI KHANDU

2016

005.76 KOS

Project Guide DR A.K.SEN

Abstract

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Modern scientific databases as well as web databases maintain large and heterogeneous data. These real-world databases contain hundreds or even thousands of relations and attributes. Traditional predefined query forms are not able to satisfy various ad-hoc queries from users on those databases. This work proposes Dynamic Query Forms (DQF), a novel database query form interface, which is able to dynamically generate query forms. The DQF captures a user's preference and ranks query form components which assists the user in making decisions. The generation of a query form is an iterative process and is guided by the user. At each iteration, the system automatically generates ranking lists of form components and the user then adds the desired form components into the query form. In this project three different ranking models are used. These are Probabilistic Ranking Model, Learning to Rank Model by using Frequent Pattern and Bayesian Network Model. A user can also fill the query form and submit queries to view the query result at each iteration. The experimental evaluation compares between these three Ranking Models and demonstrates effectiveness and efficiency of the system.

[48] PR1650 SECURITY ENHANCEMENT OF ADVANCED ENCRYPTION STANDARD(AES)USING HYBRID APPROACH

By DSOUZA,FLEVINA JONESE

2017

005.82 DSO

Project Guide MRS DAKSHATA PANCHAL

Abstract

With the fast progression of digital data exchange in electronic ways over network applications, security is much more important in data storage, communication and message transmission on public networks. During transmitting any data to a remote machine it is encrypted at the sender side for data security to perform end-end transmission and it is achieved by Cryptography (secret writing). Cryptography is the encryption process of transforming messages to make them secure and immune to attack. Advanced Encryption Standard (AES) is a symmetric encryption

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

standard recommended by the National Institute of Standards and Technology (NIST) in 2001. NIST announced that AES is a highly secure, faster and str DR A.K.SEN ong encryption algorithm. Also, it is used frequently because of its high efficiency and easy implementation. But in recent years cyber-attacks are constantly evolving, which perform attacks on AES. Possible attacks on symmetric algorithms can be Brute-force Attack, Differential Attack, Algebraic Attack and Linear Attack. In order to provide strong security in message transmission, AES algorithm with hybrid approach of Dynamic Key Generation and Dynamic S-box Generation is proposed. In hybrid approach first we add more complexity in data to increase Confusion and Diffusion in Cipher text by using Dynamic Key Generation and then by using Dynamic s-box Generation we make it difficult for attacker to do any offline analysis of static set of s-box.

[49] PR1653 A NOVEL TECHNIQUE TO INFER USER SEARCH GOALS WITH FEEDBACK SESSIONS

By MISHRA,PRITI ALOK

2017

004.019 MIS

Project Guide MRS BIDISHA ROY

Abstract

Web search engines have become a resource pool to gain information. In web search applications, queries are submitted to search engines to represent information needs of users. But there are situations in which these queries do not represent a user's specific information needs. Different users require information on different aspects when they submit the same query. Existing search engines such as Google, Yahoo return a long list of search results, ranked by their relevance to the specified query. To improve web search engine relevance, user's satisfaction of results is very important to infer user search goals. So a Novel Technique to Infer User Search Goal with Feedback Session is proposed. User Search Goals are analyzed by using a feedback session which is derived from click-through logs. Feedback sessions tell what user requires and what the user does not require. Keywords in the query will determine whether a document

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

satisfies the user's needs. The original search results will be restructured and its performance will be evaluated by using Classified Average Precision (CAP). The main objective of this technique is to improve information retrieval by generating feedback sessions and rank the results list based on the user's domain of interest. These re-ranked results will be categorization based which will be dependent on the keywords generated.

[50] PR1654 ENHANCED COOPERATIVE BAIT DETECTION SCHEME (ECBDS) AGAINST COLLABORATIVE ATTACKS IN MANETS

By GHARAT, POOJA UMESH

2017

621.3845 GHA

Project Guide MRS RAJKUMAR SHENDE

Abstract

Mobile adhoc networks (MANETs) are open medium where nodes are prone to various malicious attacks, intrusion, spying, etc. There is no centralized entity that can help in implementing security algorithms, which can overcome the attacks. There are two types of mechanism, i.e., proactive and reactive routing for route maintenance. The base protocols recommended in MANETs do not consider security at a higher level. Present protocols such as dynamic source routing (DSR) protocol and ad hoc on demand distance vector (AODV) reactive routing protocol are still not better in terms of security. Collaborative black hole attack and gray hole attacks in MANETs launched by malicious nodes are challenging to detect. Collaborative bait detection scheme (CBDS) to detect the malicious nodes launching collaborative attacks in MANETs has been proposed. This research work grants a variant in the CBDS scheme on AODV protocol by implementing a dynamic threshold on packet delivery ratio in order to reduce the false positive rate of detection of malevolent nodes. The study is conducted on network simulations using NS2 simulator by performing collaborative black hole attack and using enhanced CBDS to detect and avoid the attacks. Examination of improved performance of MANET in AODV protocol in provisions of packet delivery ratio, end to end delay and throughput is also conducted

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[51] PR1661 PREVENTION OF SESSION HIJACKING WITH MODIFIED ONE TIME COOKIES(OTC)

By SATHIYASEELAN,ANNIES MINU

2015

005.8 SAT

Project Guide MRS ANURADHA & MRS VINCY JOSEPH

Abstract

Most Transmission Control Protocol(TCP) connections use Hypertext Transfer Protocol(HTTP) to communicate; so, it becomes mandatory for every server to create a unique identifier for each and every connection. A session Identification number(ID) is a unique identifier generated by a server that is sent to a client for identifying current interaction sessions, which is stored in a cookie. A cookie is a short text file for identifying a particular client. Since cookies are transmitted over HTTP, they are visible and prone to attacks such as session hijacking. Hypertext Transfer Protocol Secure(HTTPS) is widely recommended to protect cookies, but deploying full HTTPS support can be quite challenging, especially for highly distributed applications due to performance and financial concerns. Hence, one-time cookies (OTC) are suggested as an alternative for session authentication. OTC prevents attacks, such as session hijacking, as they are temporarily stored for a particular period of time or only for a particular session. In this project, we propose a mechanism that uses OTC to prevent an attacker from gaining access to a cookie and backend server. A reverse proxy server with OTC, IP, session ID, and browser fingerprinting are used to prevent adversaries from capturing session credentials.

[52] PR1663 IMPROVISED ASPECT BASED OPINION MINING FOR PRODUCT RANKING

By BLETY BABU ALENGADAN

2018

005.74068 BLE

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Project Guide MR SHAMSUDDIN KHAN

Abstract

Due to the advancement in technology the number of fraudulent activities is increasing dramatically in every field. Fraud is a major problem in electronic payment systems. There has been the extensive and tremendous use of the Internet has led to the trust in word-of-mouth through social networking sites. A product that has the best reviews is preferred to the similar kind of product with few favorable reviews. These consumer reviews are rich in knowledge but are unstructured since they are typed by common people using these sites. Aspect/Feature based Opinion Mining is modified in this developed system to help consumers and entrepreneurs in the analysis of these user reviews by determining the strengths and weaknesses of any products. The main motto of the project is ranking the products and its important aspects which would eventually gear up a faster decision-making. The developed system includes four important phases: Pre-processing, Enhanced Aspect Identification and Opinion Word Extraction with modified Naïve Bayes model, Aspect Polarity Identification, Products and Aspects Ranking in sequence. The developed system will be beneficial for the consumers and business analysts/entrepreneurs who seek customers' online opinions for analyzing purchase decisions and market statistics and strategies

[53] PR1665 AUTHENTICATING MESSAGE IN WIRELESS SENSOR NETWORKS

By NADAR, JEBA SANGEETHA

2017

681.2 NAD

Project Guide MRS JAYASHRI MITTAL

Abstract

Message authentication is one of the most effective ways to thwart unauthorized and corrupted messages from being forwarded in wireless sensor networks

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

(WSNs). For this reason, many message authentication schemes have been developed, based on either symmetric-key cryptosystems or public-key cryptosystems. Most of them, however, have the limitations of high computational and communication overhead in addition to lack of scalability and resilience to node compromise attacks. To address these issues, a polynomial-based scheme was recently introduced. However, this scheme and its extensions all have the weakness of a built-in threshold determined by the degree of the polynomial: when the number of messages transmitted is larger than this threshold, the adversary can fully recover the polynomial. In this paper, we propose a scalable authentication scheme based on elliptic curve cryptography (ECC). While enabling intermediate nodes authentication, our proposed scheme allows any node to transmit an unlimited number of messages without suffering the threshold problem. In addition, our scheme can also provide message source privacy. Both theoretical analysis and simulation results demonstrate that our proposed scheme is more efficient than the polynomial-based approach in terms of computational and communication overhead under comparable security levels while providing message source privacy.

[54] PR1666 ESTIMATING INTRINSIC VALUE OF STOCK USING MONTE CARLO SIMULATION

By SIDDIQUI,SEHBA SHAHABUDDIN

2017

006.31 SID

Project Guide MRS VANDANA PATIL

Abstract

Stock markets for the last hundred and fifty years have been used for wealth generation. Provided people participate as an investor and invest in stocks which show strong fundamental intrinsic values. Fundamental analysis is significant from the perspective of evaluation of stock market price for long term investment. This system provides a guidance mechanism based on monte carlo simulation and discounted cash flow analysis model which is used to evaluate the list of stocks

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

under study what is their probability of becoming multi baggers, from a long term of investment perspective. The system is modeled to handle low debt companies for long term evaluation. Typically, companies having less than twenty five percent debt in their assets. The users of the system are fundamental analysts i.e. chartered accountants, chartered financial planners, chartered financial analysts, credit risk analyst, financial analyst, and portfolio managers. Comparing statistically calculated intrinsic values and current market price, the system will be able to add a robust statistical reasoning for investment decisions. This reasoning will have no human or emotional biases as there will be no human intervention involved for arriving to the final intrinsic value of stock. Monte Carlo Simulation is the best suited solution for generating random scenarios that fall in line with brownian walk motion of stock prices, as in long term any stock price follows random path.

[55] PR1667 SMART TRAFFIC CONTROLLER USING FUZZY INFERENCE SYSTEM(STCFIS)

By PATIL,RADHIKA AJIT

2015

620.001 PAT

Project Guide MRS ANURADHA

Abstract

The heavy traffic congestion problem in major cities and towns is mainly attributed to the time spent across crowded traffic junctions especially during peak hours. The traffic across junctions is controlled by signaling patterns, where the signal duration is invariably static in nature. The traffic density across either side of the road does not affect the duration of signal.

For reducing the time taken in waiting across traffic junction, the proposed work aims at simulating a dynamic traffic signal using fuzzy logic which would change the timing of the green signal with respect to the intensity of traffic and feed this to an arduino microcontroller based fuzzy inference system which decides the duration of the green signal. The intensity of traffic is computed using two methods, feature detection and image subtraction

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

methods. The feature detection method takes a lot of processing time and image subtraction method along with fuzzy logic was found more suitable in dynamically controlling the duration of traffic signals.

[56] PR1798 SENTIMENT ANALYSIS BASED ON MULTIPLE REVIEWS

By DSOUZA,STEPHINA RODNEY

2019

006.312 DSO

Project Guide DR KAVITA SONAWANE

Abstract

Sentiment Analysis can be defined as the process of analyzing online pieces of writing to determine the emotional tone they carry. With the vast growth of the social media content on the Internet in the past few years, people now express their opinion on almost anything in discussion. Finding the opinion sites and monitoring them on the web is a difficult task. Sentiment analysis is also known as “Opinion Mining” which studies people’s sentiments towards certain entities. With respect to this, Bag-of-Words (BoW) is the most popular way to model text in statistical machine learning (ML) approaches. However, the performance of BoW sometimes remains unlimited due to some fundamental deficiencies in handling the polarity shift problem and other few challenges like quality of the opinions, polarity categorization, accuracy etc. The proposed system displays the results for both the existing technique i.e. the Sentiment Analysis and the proposed Dual Sentiment Analysis (DSA), thereby comparing both the techniques and generating an accurate solution. The proposed system is executed with various classification algorithms like Naïve Bayes’, Support Vector Machine (SVM), Maximum Entropy classifier. Also, the problem of neutral comments which are generally not considered are overcome by using the split function which helps to segregate them into positive sentiment and negative sentiment, thus improving the ambiguity of the problem. The advantage of split function is that it provides a clear idea about the text/review i.e. whether it is a positive sentiment or a negative sentiment. In order to come across the above mentioned constraints and challenges, we develop a Dual Sentiment Analysis (DSA), which will not only reduce the ambiguity but also help

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

the customer in the decision making process by generating an accurate solution. This work aims to overcome the limitations of the existing system by proposing a system for maintaining the accuracy.

[57] PR1799 MEDICAL IMAGE ANALYSIS USING MACHINE LEARNING

By ALVA, MICHELLE MAXIM

2019

616.07540285 ALV

Project Guide DR KAVITA SONAWANE

Abstract

In the advent of computer aided diagnosis(CAD), analysis of medical images with the help of various feature extraction techniques is gaining a lot of interest. Various images used in the medical domain for diagnostic purposes are Computed Tomography(CT), Magnetic Resource Imaging (MRI). Positron Emission Tomography(PET), X-Ray and ultrasound images. Nowadays, a lot of research is done in the analysis of neuro-image, particularly in the MRI, modality to study the structural changes or atrophy of the brain due to the various neurological disorders. Alzheimer's disease(AD) being one such illness, affects a vast amount of elderly population worldwide. The current diagnosis of AD which is done manually by a clinician using neuro-images can sometimes be inaccurate. Although image processing and machine learning facilities feature extraction and diagnosis of AD respectively, there are still many challenges that exist in the current systems. In this work volumetric and textural analysis of brain tissues such as grey matter (GM) and white matter (WM) is performed using the T1 weighted sMRI images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset and the combined effect of features obtained from various approaches is studied. This work explores different feature extraction techniques including whole-brain and ROI based analysis such as Voxel-based Morphometry (VBM) to study changes in volume of the brain, and histogram bins-approach and Grey-level Co-occurrence Matrix(GLCM) approach for textural analysis. Based on the results obtained, a hybrid feature vector(HFV) is proposed which has the

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

collective strength of the standalone feature extraction techniques with the help of which the system will be able to classify the subjects as Alzheimer's disease(AD) and Normal Controls(NC) using various machine learning approaches.

[58] PR1805 Crop Yield Prediction using a Fusion of Landsat and Sentinel Data with Meteorological Parameters

By Monisha

2021

338.1MON

Project Guide Dr. Vaishali Jadhav

Abstract

Crop yield prediction is an important agriculture problem. Most of the existing work focused on predicting the yield of Rice and other crops like Maize, Cotton, Sugarcane, etc. Proposed research work focuses on wheat and jowar crops which are of the same season i.e. Rabi season. This proposed work used the combination of Landsat and Sentinel data with meteorological parameters. This paper aims to predict the yield of wheat crops. The satellite images of Maharashtra state were collected for this proposed work. The various growing stages of wheat are considered for the yield prediction model. The same dates of Landsat and Sentinel images for various growing months are collected. Then these images were preprocessed (atmospheric correction) and the Landsat and Sentinel images were combined to find the vegetation index of the crop (NDVI). This vegetation portion was mainly used to find the yield of the crop. Then Normalized Difference Vegetation Index (NDVI) value is combined with meteorological parameters such as rainfall, temperature, humidity and find the yield using the ensemble learning technique. This is the combination of Random forest and Ada-boost algorithm. This technique is mainly used to enhance the accuracy of the yield prediction model. Ensemble learning is mostly used in complex real time problems to reduce the variance and minimize bias. This proposed work compares the yield prediction

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

model of wheat and jowar and it also examines that which crop will give better yield in rabi season.

[59] PR2061 Location Prediction on Twitter

By Chetali Surti

2021

006.3 SUR

Project Guide Ms. Pradnya Rane

Abstract

Online social networks have gained a lot of popularity over the last decade, especially. Twitter. In every second thousands of tweets are tweeted on Twitter which produces massive amounts of data every day in the whole world. With the tweet, peoples frequently share their current location by means of check-ins point of interest that is recognized as geo-tagged tweet. On the other hand, most of the people do not mention their locations with the tweets. Location prediction is advantageous for a number of fields such as advertising, rescue operations during a disaster and recommendation, etc. Location prediction from the text of the tweet is a difficult task as it may contains many spelling mistakes and non-standard language. Hence location prediction on Twitter has attracted many researchers now days. The proposed system aims to use hybrid model taking advantage of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) for Location Prediction on Twitter using text of tweet as input.

[60] PR2063 SECURING FILES USING IMAGE ENCRYPTION AND DECRYPTION

By DSOUZA/CLETUS ANDREW AGNELLO

2020

005.82 DSO

Project Guide KAVITA SONAWANE

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

Abstract

Encryption is the process of encoding a text, image or file in order to secure it. In the world of social media there is a vast growth in online content and ways of communication. Online contents may be in various forms like text, image, audio, video etc. It is a need of the hour for several system developers to protect the user contents in any forms. There are various commonly used encryption algorithms used by researchers for file security like DES, AES, Blowfish, Chaos Approach. Based on review of literature it is found that each of these algorithms is giving varying performance as context changes and there are some constraints observed among few of these algorithms. The proposed work aims at addressing a few of the challenges as issues and constraint of existing methods by identifying problems areas and devising a solution. Based on the recent study, the primary focus of this work would be Chaos and AES. Actually, it uses a modified chaos which has proved better in terms of results by creating a high level of confusion while encrypting the file to be secured. The unique idea is to introduce Seed key which is equal to the MAC address of PC and Bluetooth address of android phone which gives full control to the user and intruder would be confused from where the values are being bought in to the picture, the use of such key will promote more randomness. After seed key obtained proceeding steps include containing image key by chaos and then that would encrypt by the Rijndael algorithm which in turn will complete the dual encryption.

So the proposed system will try to contribute better encryption for image security. Performance of the proposed algorithm will be evaluated using key sensitivity, NPCR and UACI, histogram analysis in order to test the system w.r.t to previous works.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

[61] PR2064 Handling Class Imbalance in Fraud Detection using Machine Learning Techniques

By Reshma George

2020

006.31GEO

Project Guide Ms. Bidisha Roy

Abstract

A disproportion between the fraudulent and the non-fraudulent class is generally termed as class imbalance. In class imbalance problems the instances of the positive class (minority) are much smaller than the number of instances of the negative class (majority). The non-fraudulent class being more dominant and the fraudulent class being relatively rare the latter would be considered as outliers thereby resulting into a misclassification of the minority class. Hence the results of imbalanced classes are not always accurate. The applications of class imbalance ranges from fraud detection, anomaly detection, oil spillage detection, network intrusion to medical diagnosis. In all these scenarios we have cases where we have two different kinds of labels and one of them is highly skewed. The proposed model aims at using the sampling techniques like Random under sampling and Random oversampling to sample the dataset after which the balanced dataset will be given to the classifier and the results will be then compared and evaluated using the performance evaluation metrics.

[62] PR2203 Face Mask Detection Using Machine Learning

By Shruti Mishra

2023

006.31 MIS

Project Guide Dr. Vaishali Jadhav

Abstract

Coronavirus disease 2019 has affected the world seriously. Coronavirus belongs.

ST. FRANCIS INSTITUTE OF TECHNOLOGY
(ENGINEERING COLLEGE)

Mumbai 400103

www.sfit.ac.in

to the class of viruses, A coronavirus identified in 2019, SARS-CoV-2, has caused a pandemic of respiratory illness, called COVID-19. COVID-19 is the disease caused by SARS-CoV-2, the coronavirus that emerged in December 2019. COVID-19 can be severe, and has caused millions of deaths around the world as well as lasting health problems in some who have survived the illness. Coronavirus spreads through droplets and virus particles released into the air when an infected person breathes, talks, laughs, sings, coughs or sneezes. After sneezing or coughing the large droplets that fall into the ground causes the viruses to accumulate at one place and this leads to transfer of disease from person to person. Wearance of mask prevents the spread of such infections and disease. But achieving this practice manually, it is a very tedious task. There-fore, this demands the existence of automated face mask detection system, that identifies automatically whether the person has wore a mask or not. The detector system should be viable and has to deployable in public so as to curb the spread of disease thereby making the public to mandatorily wear the mask. In this paper we aim to perform a comparative analysis of various sophisticated im-age classifiers based on deep learning, in terms of vital metrics of performance to identify the effective deep learning-based model for face mask detection.