



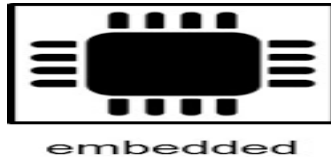
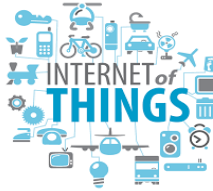
IGEEKS Technologies

Bridging Technology.

For: - B. E | B. Tech | M. E | M. Tech | MCA | BCA | Diploma | MS | M. Sc |

IEEE

REAL TIME PROJECTS & TRAINING GUIDE
SOFTWARE & EMBEDDED



FINAL YEAR PROJECTS

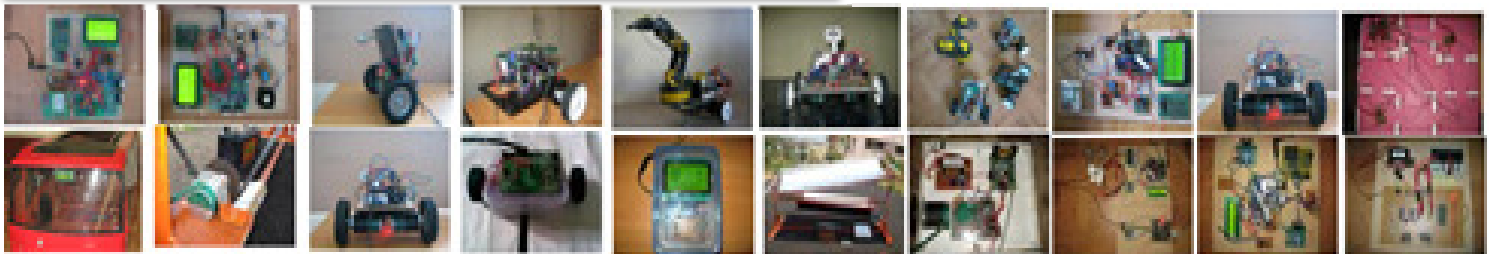
IEEE & Application Projects

BE, Diploma, BSc
M Tech, MCA, BCA

CS | Mechanical
E & C | Electrical

www.makefinalyearproject.com

7019280372/9590544567



PYTHON ML, DATA SCIENCE, AI PROJECT ABSTRACTS FOR ACADEMIC YEAR 2019 - 2020

#19, MN Complex, 2nd Cross, Sampige Main Road, Malleswaram, Bangalore - 560003

Call Us: 9590544567 / 7019280372 www.makefinalyearproject.com

www.igeeekstechnologies.com Land Mark: Opposite Joyalukkas Gold Showroom, Near to Mantri Mall

[1] A Compendium for Prediction of Accidents Severity in India by Data mining Approach.

Road accident analysis plays an important role in transportation system. This project shows a survey of road accident analysis methods in data mining. In the data mining there is no. of techniques available for clustering and classification, from those techniques k-mean, association rule, SVM, Weka tool was used in previously research for road accident analysis. In our daily life there are no. of accident increases and it is big problem to us because no. of people death and injured for that improve the road transportation system is needed. In this research self-organization map (SOM) is used for find a no. of pattern to analysis the road accident data which help to find prediction of accident reasons and improve the accuracy of analysis compare to k-means clustering algorithm. With the help of SOM, clusters are created and analyze them. Self-Organizing map method is based on neural network, it is used as an unsupervised learning method. It will help to improve analysis accuracy.

[2] Predictive Analytics Model to Diagnose Breast Cancer Tissues using SVM classifier.

The cancer is the most dangerous diseases in the world, its mainly effective for women. So, our prime target must be curing the cancer through scientific investigation and the second main target should be early detection of cancer because the early detection of cancer can be helpful for remove the cancer completed. After reviewed 41 projects we found that several techniques are available for cancer detection. In this project we proposed Deep Learning algorithm neural network for diagnosed breast cancer using Wisconsin Breast Cancer database. The project shows how we can use deep learning technology for diagnosis breast cancer using UCI Dataset. Because deep learning techniques almost used for high task objective Computer Vision, Image processing, Medical Diagnosis, Neural Language Processing. But in this

project, we are applying deep learning technology on the Wisconsin Breast Cancer Database and we have seen that is very beneficial for us for diagnosis breast cancer with accuracy 99.67%. This project is divided in three parts first we have collect dataset and applied pre-processing algorithm for scaled and filter data then we have split dataset in training and testing purpose and generate some graph for visualization data. In last implement model on training dataset and achieved accuracy 99.67%. So, we have seen deep learning technology is a good way for diagnosis breast cancer with Wisconsin Breast Dataset. This database provides 569 rows and 30 features in the dataset. In this project we have used 11 features for diagnosis breast cancer that we have got after pre-processing. But before train model we have applied some pre-processing algorithm like Label Encoder, Normalizer and Standard Scaler for scaled dataset then applied model and achieved accuracy. In this project we also compare deep learning algorithm with other machine learning and seen our proposed system is proved best from others machine learning algorithm.

[3] Job Satisfaction and Employee Turnover in Organizations using Machine Learning Techniques: A Firm Level Perspective

An analysis of a unique data set of over 970,000 Curriculum Vitas (CVs), revealed a mild inverse relation between job satisfaction and the average employment period. We found that while higher compensation packages often lead to higher job satisfaction, this does not translate into longer employment period. Moreover, our findings indicate an opposite and surprising pattern, in which as salary and satisfaction levels increase, employment period per job decreases. This counter-intuitive effect was found to be strongest in global high technology companies.

[4] A Novelistic Supervised Learning Approach in Financial Revenue Analysis for constructing a Feasible Model to predict the Donors for Charity

Private giving represents more than three fourths of all U.S. charitable donations, about 2% of total Gross Domestic Product (GDP). Private giving is a significant factor in funding the nonprofit sector of the U.S. economy, which accounts for more than 10% of total GDP. Despite the abundance of data available through tax forms and other sources, it is unclear which factors influence private donation, and a reliable predictive mechanism remains elusive. This work aims to develop predictive models to accurately estimate future charitable giving based on a set of potentially influential factors. We have selected several factors, including unemployment rate, household income, poverty level, population, sex, age, ethnicity, education level, and number of vehicles per household. This work sheds light on the relationship between donation and these variables. We use Stepwise Regression to identify the most influential variables among the available variables, based on which predictive models are developed. Multiple Linear Regression (MLR) and machine learning techniques, including Artificial Neural Networks (ANN) and Support Vector Regression (SVR) are used to develop the predictive models. The results suggest that population, education level, and the amount of charitable giving in the previous year are the most significant, independent variables. We propose three predictive models (MLR, ANN, and SVR) and validate them using 10-fold cross-validation method, then evaluate the performance using 9 different measuring criteria. All three models are capable of predicting the amount of future donations in a given region with good accuracy. Based on the evaluation criteria, using a test data set, ANN outperforms SVR and MLR in predicting the amount of charitable giving in the following year.

[5] Machine Learning Techniques for Human Heart Diseases Prediction with Claims Data

Data mining is an advanced technology, which is the process of discovering actionable information from large set of data, which is used to analyze large volumes of data and extracts patterns that can be converted to useful knowledge. Medical data mining has a great potential for exploring the hidden patterns in the data sets of medical domain. These patterns can be utilized to do clinical diagnosis. These data need to be collected in a standardized form. From the medical profiles fourteen attributes are extracted such as age, sex, blood pressure and blood sugar etc. can predict the likelihood of patient getting heart disease. These attributes are fed in to K-means algorithms, MAFLA algorithm and Decision tree classification in heart disease prediction, applying the data mining technique to heart disease treatment; it can provide as reliable performance as that achieved in diagnosing heart disease. By this medical industries could offer better diagnosis and treatment of the patient to attain a good quality of services. The main advantages of this project are: early detection of heart disease and its diagnosis correctly on time and providing treatment with affordable cost.

[6] Two-fold Computation Techniques for Twitter Sentimental Analysis using NLP approach.

Growth in the area of opinion mining and sentiment analysis has been rapid and aims to explore the opinions or text present on different platforms of social media through machine-learning techniques with sentiment, subjectivity analysis or polarity calculations. Despite the use of various machine-learning techniques and tools for sentiment analysis during elections, there is a dire need for a state-of-the-art approach. To deal with these challenges, the contribution of this project includes the adoption of a hybrid approach that involves a sentiment analyzer that includes machine learning. Moreover, this project also provides a comparison of techniques of sentiment analysis in the

analysis of political views by applying supervised machine-learning algorithms such as Naïve Bayes and support vector machines (SVM).

[7] Dynamic Reconstruction of Vehicle Detection alone with the lane using a High Resolution Images by Data Mining Approach.

This project presents a high performance vision-based system with a single static camera for traffic surveillance, for moving vehicle detection with occlusion handling, tracking, counting, and One Class Support Vector Machine (OC-SVM) classification. In this approach, moving objects are first segmented from the background using the adaptive Gaussian Mixture Model (GMM). After that, several geometric features are extracted, such as vehicle area, height, width, centroid, and bounding box. As occlusion is present, an algorithm was implemented to reduce it. The tracking is performed with adaptive Kalman filter. Finally, the selected geometric features: estimated area, height, and width are used by different classifiers in order to sort vehicles into three classes: small, midsize, and large. Extensive experimental results in eight real traffic videos with more than 4000 ground truth vehicles have shown that the improved system can run in real time under an occlusion index of 0.312 and classify vehicles with a global detection rate or recall, precision, and F-measure of up to 98.190%, and an F-measure of up to 99.051% for midsize vehicles.

[8] A Novelistic Approach to Analyse Weather Conditions and its Prediction using Deep learning techniques.

Due to the impact of weather forecasting on global human life, and to better reflect the current trend of weather changes, it is necessary to conduct research about the prediction of precipitation and provide timely and complete precipitation information for climate prediction and early warning decisions to avoid serious meteorological disasters. For the precipitation prediction problem in the era of climate big data, we propose a new method based on

deep learning. In this project, we will apply deep belief networks in weather precipitation forecasting. Deep belief networks transform the feature representation of data in the original space into a new feature space, with semantic features to improve the predictive performance. The experimental results show, compared with other forecasting methods, the feasibility of deep belief networks in the field of weather forecasting.

[9] Rainfall prediction using Machine learning techniques

Rainfall is one of the major source of freshwater for all the organism around the world. Rainfall prediction model provides the information regarding various climatological variables on the amount of rainfall. In recent days, Deep Learning enabled the self-learning data labels which allows to create a data-driven model for a time series dataset. It allows to make the anomaly/change detection from the time series data and also predicts the future event's data with respect to the events occurred in the past. This project deals with obtaining models of the rainfall precipitation by using Deep Learning Architectures (LSTM and ConvNet) and determining the better architecture with RMSE of LSTM as 2.55 and RMSE of ConvNet as 2.44 claiming that for any time series dataset, Deep Learning models will be effective and efficient for the modellers.

[10] Cyber Threat Detection and Analysis using Machine Learning Approach

A threat can be anything that causes potential damage to the network system. These threats can turn out to be a attack to the system. Threat may occur in any forms like viruses, outright attack, and phishing attack from hackers to gain information. Such attacks put a user's system and also business system at risks. Cyber security aims at the protection of system from attacks like unauthorized network access, intrusions attacks etc. This project presents a novel architecture model based on machine learning for the prediction of Cyber security malware that requires execution in a sandbox environment. In order

to prevent the attackers from infiltrating the system Machine Learning approach is adapted.

[11] Resolution of Word Sense Disambiguation using NLP approach and Text Analytics.

Word Sense Disambiguation is an open problem in Natural Language Processing which is particularly challenging and useful in the unsupervised setting where all the words in any given text need to be disambiguated without using any labeled data. Typically WSD systems use the sentence or a small window of words around the target word as the context for disambiguation because their computational complexity scales exponentially with the size of the context. In this project, we leverage the formalism of topic model to design a WSD system that scales linearly with the number of words in the context. As a result, our system is able to utilize the whole document as the context for a word to be disambiguated. The proposed method is a variant of Latent Dirichlet Allocation in which the topic proportions for a document are replaced by synset proportions. We further utilize the information in the WordNet by assigning a non-uniform prior to synset distribution over words and a logistic-normal prior for document distribution over synsets. We evaluate the proposed method on Senseval-2, Senseval-3, SemEval-2007, SemEval2013 and SemEval-2015 English All-Word WSD datasets and show that it outperforms the state-of-the-art unsupervised knowledge-based WSD system by a significant margin.

[12] Analysis and Prediction of Chronic Kidney Disease (CND) using Data Mining Approach.

Chronic Kidney Disease is a serious lifelong condition that induced by either kidney pathology or reduced kidney functions. Early prediction and proper treatments can possibly stop, or slow the progression of this chronic disease to end-stage, where dialysis or kidney transplantation is the only way to save patient's life. In this work, we examine the ability of several machine-learning

methods for early prediction of Chronic Kidney Disease. This matter has been studied widely; however, we are supporting our methodology by the use of predictive analytics, in which we examine the relationship in between data parameters as well as with the target class attribute. Predictive analytics enables us to introduce the optimal subset of parameters to feed machine learning to build a set of predictive models. This work starts with 24 parameters in addition to the class attribute, and ends up by 30 % of them as ideal sub set to predict Chronic Kidney Disease. A total of 4 machine learning based classifiers have been evaluated within a supervised learning setting, achieving highest performance outcomes of AUC 0.995, sensitivity 0.9897, and specificity 1. The experimental procedure concludes that advances in machine learning, with assist of predictive analytics, represent a promising setting by which to recognize intelligent solutions, which in turn prove the ability of predication in the kidney disease domain and beyond.

[13] Early detection and Analysis of Diabetes in Humans using ML approach

Healthcare domain is a very prominent research field with rapid technological advancement and increasing data day by day. In order to deal with large volume of healthcare data we need Big Data Analytics which is an emerging approach in Healthcare domain. Millions of patients seek treatments around the globe with various procedure. Analyzing the trends in treatment of patients for diagnosis of a particular disease will help in making informed and efficient decisions to improve the overall quality of healthcare. Machine Learning is a very promising approach which helps in early diagnosis of disease and might help the practitioners in decision making for diagnosis. This project aims at building a classifier model using WEKA tool to predict diabetes disease by employing Naive Bayes, Support Vector Machine, Random Forest and Simple CART algorithm. The research hopes to recommend the best algorithm based on efficient performance result for the prediction of diabetes disease.

Experimental results of each algorithm used on the dataset was evaluated. It is observed that Support Vector Machine performed best in prediction of the disease having maximum accuracy.

[14] Email Spam Filter analysis and its classification

The upsurge in the volume of unwanted emails called spam has created an intense need for the development of more dependable and robust antis spam filters. Machine learning methods of recent are being used to successfully detect and filter spam emails. We present a systematic review of some of the popular machine learning based email spam filtering approaches. Our review covers survey of the important concepts, attempts, efficiency, and the research trend in spam filtering. The preliminary discussion in the work background examines the applications of machine learning techniques to the email spam filtering process of the leading internet service providers (ISPs) like Gmail, Yahoo and Outlook emails spam filters. Discussion on general email spam filtering process, and the various efforts by different researchers in combating spam through the use machine learning techniques was done. Our review compares the strengths and drawbacks of existing machine learning approaches and the open research problems in spam filtering. We recommended deep leaning and deep adversarial learning as the future techniques that can effectively handle the menace of spam emails.

[15] Live Facial Expression Recognition using Artificial Neural Network

Automatic recognition of human affects has become more interesting and challenging problem in artificial intelligence, human-computer interaction and computer vision fields. Facial Expression (FE) is the one of the most significant features to recognize the emotion of human in daily human interaction. FE Recognition (FER) has received important interest from psychologists and computer scientists for the applications of health care assessment, human affect analysis, and human computer interaction. Human express their

emotions in a number of ways including body gesture, word, vocal and facial expressions. Expression is the important channel to convey emotion information of different people because face can express mainly human emotion. This project works on the current research works related to facial expression recognition. It attends to explored details of the facial datasets, feature extraction methods, etc.

[16] Black Friday: Market Trade analysis and Prediction of sales in the market

During the Black Friday sale, all the retail shops are crowded. Most products are marked down with discounts and customers rush in to buy the products. It is difficult for customers to buy the products even with a solid plan. But, the shop owners face even more difficulty on controlling the crowd with limited staff and in targeting prospective customers. Several techniques have been employed to tackle this problem, but they are not that successful. A prediction model is a technique that has proved promising in solving the problem. This work focuses on the field of prediction models to develop an accurate and efficient algorithm to analyze the customer spending in the past and output the future spending of the customers with same features. In this work, different machine learning techniques such as regression and neural network to develop a prediction model are implemented and a comparison is done based on their performance and accuracy of prediction. These techniques are implemented using different algorithms and on different platforms to find the best predication. We implemented seven different machine learning algorithms. Further, this work discusses the data pre-processing and visualization techniques employed to attain the optimal results.

[17] Market analysis and Crop yield prediction using Machine learning techniques.

Agriculture is the most important sector that influences the economy of India. It contributes to 18% of India's Gross Domestic Product (GDP) and gives employment to 50% of the population of India. People of India are practicing Agriculture for years but the results are never satisfying due to various factors that affect the crop yield. To fulfill the needs of around 1.2 billion people, it is very important to have a good yield of crops. Due to factors like soil type, precipitation, seed quality, lack of technical facilities etc the crop yield is directly influenced. Hence, new technologies are necessary for satisfying the growing need and farmers must work smartly by opting new technologies rather than going for trivial methods. This project focuses on implementing crop yield prediction system by using Data Mining techniques by doing analysis on crops dataset.

[18] Credit card fraud detection using Machine learning techniques.

Credit card fraud events take place frequently and then result in huge financial losses. Criminals can use some technologies such as Trojan or Phishing to steal the information of other people's credit cards. Therefore, an effective fraud detection method is important since it can identify a fraud in time when a criminal uses a stolen card to consume. One method is to make full use of the historical transaction data including normal transactions and fraud ones to obtain normal/fraud behavior features based on machine learning techniques, and then utilize these features to check if a transaction is fraud or not. In this paper, two kinds of random forests are used to train the behavior features of normal and abnormal transactions. We make a comparison of the two random forests which are different in their base classifiers, and analyze their performance on credit fraud detection.

[19] Student performance prediction

Data Mining methods are applied on educational data with the intent of enhancing teaching methods, improving quality of teaching, identifying weak students, identify factors that influence Student's academic performance. This utilization of data mining methods to elevate quality of education, identifying students who need improvement is termed as educational data mining. EDM has become a major research interest for many researchers. The primary function of educational data mining is prediction of student's academic performance. [1] Predicting student's academic performance helps in identifying a number of things like students who are likely to drop out, students who are weak and needs improvement, students who are good in academics but lately deteriorated. The intent of this paper is to determine factors that can influence a student's academic performance.

[20] Phishing attacks detection and classification.

Phishing is a form of cybercrime where an attacker imitates a real person / institution by promoting them as an official person or entity through e-mail or other communication mediums. In this type of cyber attack, the attacker sends malicious links or attachments through phishing e-mails that can perform various functions, including capturing the login credentials or account information of the victim. These e-mails harm victims because of money loss and identity theft. In this study, a software called "Anti Phishing Simulator" was developed, giving information about the detection problem of phishing and how to detect phishing emails. With this software, phishing and spam mails are detected by examining mail contents.

[21] Movie Sentiment analysis using ML techniques

Sentiment Analysis is a new subject in Research and is useful in many other fields. In Modern World, A huge amount of textual data is collected using surveys, comments, and reviews over the web. All of the collected data is used to improve products and services provided by both private organizations and governments around the world. This Paper includes sentiment analysis of movie reviews using feature-based opinion mining and supervised machine learning. In this paper, the main focus is to determine the polarity of reviews using nouns, verbs, and adjectives as opinion words. Reviews will be Classified into two different categories positive and negative. Reviews of Open Movie Database is used as source data set and Natural Language Processing Toolkit for Part of Speech Tagging. This paper also contains some facts about the classification of data on basis of polarity

[22] A personalized Movie recommendation using a Hybrid approach of structural balance theory

In the spread of information, how to quickly find one's favorite movie in a large number of movies become a very important issue. Personalized recommendation system can play an important role especially when the user has no clear target movie. In this paper, we design and implement a movie recommendation system prototype combined with the actual needs of movie recommendation through researching of KNN algorithm and collaborative filtering algorithm. Then we give a detailed principle and architecture of JAVAEE system relational database model. Finally, the test results showed that the system has a good recommendation effect.

[23] Prediction of the Diabetes readmission of the patients.

Hospital readmissions within 30 days after discharge are costly and it has been a prior for researchers to identify patients at risk of early readmission. Most of

the reported hospital readmission prediction models have been built with historical data and thus can outdate over time. In this work, a self-adaptive 30-day diabetic hospital readmission prediction model has been developed. A diabetic inpatient encounter data stream was used to train the self-adaptive models based on incremental learning algorithms. The result indicated that the model can automatically adapt to the newly arrived data. The best model achieved an average AUC of 0.655 ± 0.078 , which is consistent with static models built with the same dataset.



IGEEKS Technologies

Bridging Technology.

Head Office:

IGEEKS Technologies
No:19, MN Complex, 2nd Cross,
Sampige Main Road, Malleswaram,
Bangalore Karnataka (560003) India.
Above HOP Salon,
Opp. Joyalukkas, Malleswaram, Land
mark : Near to Mantri Mall, Malleswaram
Bangalore.

Email: nanduigeeks2010@gmail.com,
nandu@igeekstechnologies.com

Office Phone:
9590544567 / 7019280372

Contact Person:

Mr. Nandu Y,

Director-Projects,

Mobile: 9590544567,7019280372

E-mail: nandu@igeekstechnologies.com
nanduigeeks2010@gmail.com

**No.1 Rated company in Bangalore for all
software courses and Final Year Projects**

Igeeks Technologies
(Malleswaram)

★★★★★

Excellent	78.9 %
Very Good	10.5 %
Good	10.6 %
Average	0 %
Poor	0 %

Please rate on
JustdialTM

JD App available on:

As on Jun 2015

RAJAJINAGAR:

#531, 63rd Cross,
12th Main, after sevabhai hospital,
5th Block, Rajajinagar,
Bangalore-10.
Landmark: Near Bashyam circle.

JAYANAGAR:

#65, 'Bhagyadeep', 8th 'B' Main, 27th Cross,
Jayanagar 3rd Block (Next to Pizza
Hut), Bangalore 560011.

More than 13 years' experience in IEEE Final Year Project Center, IGEEKS Technologies Supports you in Java, IOT, Python, Bigdata Hadoop, Machine Learning, Data Mining, Networking, Embedded, VLSI, MATLAB, Power Electronics, Power System Technologies.

For Titles and Abstracts visit our website www.makefinalyearproject.com