SRK INSTITUTE OF TECHNOLOGY

ENIKEPADU, VIJAYAWADA

AI SPYDER CLUB

Artificial Intelligence is a tool, not a threat

June, 2019

Faculty Advisors
Dr. D. Haritha
N. Sudhakar
P. Rani

Co-Ordinators
E. Jahnavi
Siva Sai Babu
Sk. Sabiya
Kundan
T. Nikhitha
Bala Kamal

PROJECT ABSTRACTS

RUMOUR DETECTION ON SOCIAL MEDIA

16X41F0001 | A. Tara Siva Durga

Automatically identifying rumours from online social media especially micro-blogging websites is technically very challenging and an important research. Experiment conclusion show that our emerging rumour identification for social media with hot topic detection work well both in news data set and twitter dataset, and linking the hot topic detection with the rumour detection is achievable to finish real-time rumour identification.

The recommended model is based on Recurrent Neural Networks(RNN). To study the hidden representation that capture the variation of contextual information of related posts over time. Experimental conclusion on datasets from real-world microblog platforms demonstrate that RNN-based method detects rumours more immediately and exactly than existing approaches including the leading online rumour debunking services

CLOUD CLASSIFICATION AND WEATHER PREDICTION USING DIGITAL WEATHER IMAGES

16X41F0006 | D.DHATHATHREYA

Weather forecasting is the process of recording the parameters of weather like wind direction, wind speed, humidity, rain fall, temperature etc. From these parameters the prediction of rain fall estimation can be done using different types of digital weather images.

Convolution Neural Network (CNN) can be applied on these parameters in order to predict the future rainfall. In "Cloud classification and weather prediction using digital weather images" technique, these parameters say humidity, temperature and rainfall will be predicted with respect to temperature.

In this project, different models which are used in the past for weather forecasting are discussed. This project is focusing on three parts, first is different models which were used in weather forecasting, second part is introducing new wireless kit use for weather forecasting, and third part includes the Convolution Neural Network(CNN) can be applied on different digital images for weather forecasting.

PREDICTION OF LIVER DISEASE BASED ON DEEP LEARNING

16X41F0008 | J. Prathyusha

Diagnosis of liver disease at a preliminary stage is important for better treatment. It is a very challenging task for medical researchers to predict the disease in the early stages owing to subtle symptoms. Often the symptoms become apparent when it is too late. To overcome this issue, this project aims to improve liver disease diagnosis using machine learning approaches.

The main objective of this research is to use classification algorithms to identify the liver patients from healthy individuals. In this study, THREE classification algorithms Logistic Regression, Decision tree and Recurrent neural networks (RNN) have been considered for comparing their performance based on the liver patient data. Further, the model with the highest accuracy is implemented. The dataset used is The Indian Liver Patient Dataset (ILPD) which was selected from UCI Machine learning repository for this project.

FACE RECOGNITION BASEDATTENDANCE SYSTEM USING MACHINE LEARNING

16X41F0009 | J. RAM BABU

Attendance is an important part of daily classroom ascertainment for the teacher for his or her smooth running of class. At the beginning and ending of the class usually attendance is checked by the teacher, but the manual attendance system may leads to appear that a teacher may miss someone or some students may answer multiple times. Face recognition-based attendance system is a problem of recognizing face for taking attendance by using face recognition technology based on high definition monitor video and other information technology. Now days, **Machine Learning** has been highly explored for computer vision applications. So, we use the concept of machine learning in Face – recognition for automatic attendance systems. In this machine learning we perform the face recognition and face detection algorithms, by giving input to a computer system the ability of finding and recognizing human faces fast and precisely in images or videos.

FRAUD AND ABUSE DETECTION FOR INSURANCE COMPANIES

16X41F0021 | P.TIRUMALARAO

In the recent days information is wealth and showing the false information which turns as a fraudulent crime now a days to generate the income. In order to predict such frauds that occurs in the insurance claims are not well aware to the public. This Project demonstrates machine learning tools and techniques to discover such fraud in large sets of insurance claimed data in insurance.

Such kind of insurance claim may occur by showing the fraudulent information and claiming the insurance policy .This research gives plan machine learning based fraud detection technique for insurance companies using the random forest detection analysis is proposed to find fraudulent records in auto insurance and all other insurance types.

ADVERTISEMENT CLICK FRAUD DETECTION

16X41F0022 | P. ANUSHA

On-line advertisement has become one of the main important earning method that to support Internet sites. Given that large amount of money are involved in on-line advertisement. Cheating persons are unfortunately trying to gain an unfair advantage of these on-line advertisements. Click-fraud attacks are one instance of such malicious behaviour, where software imitates a human clicking on an advertisement link.

Basically, online advertisers used the Cost-Per-Impression (CPI) model to charge for advertisements. The CPI is often measured in terms of the cost of one thousand impressions of the advertisement. Such advertisements are videos, images and links. Recently, search engines such as Google have given rise to the Pay-Per-Click (PPC) model for online advertising. In this type of package, advocate pay a certain load of money to the publish for every click on their ad (which leads to the advertiser's website). This model has however led to the rise of click fraud where the publisher, his employees or a leased botnet of computers fraudulently click on the ads in order to develop credit for the journalist in an invisible manner. Another form of click fraud happens when competitors of an advertiser click on his ads in order to deplete his daily budget. In another scenario, some advertisers or publishers may cooperate against a single publisher by clicking on his ads, in order to force the broker to exclude the publisher from the ad campaigns.

we perform analysis on fraud detection which happens in online advertising modules, together with standard threshold-based methods and machine learning algorithms.

FAKE NEWS DETECTION ON SOCIAL MEDIA USING MACHINE LEARNING

16X41F0023|S. SAI TARUN

Now a day's Fake news spreading on social media has recently become captive research that is attracting tremendous Attention to Everyone. Mainly spreading Fake news on social media to spoil someone's personal life.

Identifying Fake news on social media presents special and unique characteristics and challenges that make existing machine learning detection algorithms from traditional news media ineffective or not applicable.

In this, we present a comprehensive review of detecting fake news on social media, including fake news characterizations on psychology and social theories here we have some machine learning algorithms to detect or Identify fake new.

STUDENT PERFORMANCE PREDICTION.

16X41F0025 | U. HEMA GAYATHRI

The ability to predict the academic performance based on the behavioural characteristics and regular performance of students is vital in today's competitive world. Like other domains, educational domain also produce vast amount of data for analysis. To enhance the quality of education system, Student Performance Analysis plays an important role for decision support. It helps the teacher to identify students who need special attention and allow the teacher to provide appropriate guidance.

The data set taken contains several attributes or features are considered, like Roll no, Name, Gender, Attendance, Graduation, Year, University, Unit tests, Mid & Sem Marks, Board marks, Grade, GPA, Locality, Result, Day to day performance etc. On basis of these attributes student performance can be predicted by using advanced analytical techniques.

In this project, Machine learning techniques can be utilized for student's performance prediction. Because the vast amount of data can't be easily predicted by human where as machine learning techniques can handle it easily. This project showcases the power of Machine Learning techniques in variety of domains.

PREDICTION OF HEART ATTACK USING SPARK AND DEEP LEARNING

16X41F0026 | V. HIMA ANNAPURNA

Prediction of a heart attack plays a crucial role since it is one of the leading causes of sudden death. The healthcare industry produces large amounts of health care data daily that can be used to extract information for predicting heart attack that can happen to a patient in future while using the treatment history and health status. This hidden information in the health care data will be later used for affective decision making for patient's health.

Researchers are focussing on developing software with the help of machine learning algorithms which can help doctors to take decision about patient's health regarding both prediction and diagnosing of heart disease.

The main objective of this project is predicting the heart disease of a patient using machine learning algorithms. Comparative study of the various performances of machine learning algorithms will be done through graphical representation of the results.

The objective and scope of this project is to predict the chances of getting heart attack before it occurs, then early treatment can be given to the patients which can reduce risk to the life and save life of patients and reduce cost of the treatment .For this, a probabilistic modelling and deep Learning approach using Long Short-Term Memory (LSTM) recurrent neural network and two convolutional neural networks for prediction of heart attack are being used.

PLANT DISEASE DETECTION USING CONVOLUTIONAL NEURAL NETWORKS.

16X41F0027 | V. PRAVALLIKA

Agriculture is one of the field on which countries economy depends. The resources of products like food, paper and cosmetics are mainly from the trees. The quality of the above products is mainly depend on the quality of supplied resources. These product resources which in turn relay on the plants growth and yield they get. There is a necessity of supplying good quality resources in order to produce quality products. Detection of various diseases at early stages of plants plays an important role in the growth of plants. If the disease is identified at the early stage one can take the necessary action to reduce the effect of diseases on the growth of the plants which in turn helps in producing good quality products.

Convolution neural network (CNN) is used in image recognition and processing that is specifically design to process pixel data. The project aims to develop CNN based plant disease identification system which will certainly helpful to reduce the effect of diseases on the plants. The ability of CNN is that it can extract the hidden features of plant images through several layers, which is not possible through the convectional image feature extraction methods.

SPAM SMS FILTERING USING MACHINE LEARNING.

16X41F0012 | K. PAVITHRA

Short Message Service(SMS) is one of the well-known communication services in which a message is send electronically. The lessening in the cost of SMS benefits by telecom organisation has prompted the expanded utilisation of SMS. Here rise drawn in attacker, which have brought about SMS Spam problem. Spam messages include announcement, free services, promotions, awards, etc. People are using everywhere of mobile phone devices is spread everywhere as they give a vast variety of services by reducing the cost of services. Short Message Service(SMS) is one of the broadly utilized communication services. In any case, this has move growth in mobile phones attacks like SMS Spam. In this problem, preliminary results are mentioned or explained here in based on Singapore based publically available datasets. This problem is extra increase using multiple background datasets. In the proposed system we are going to use machine learning techniques for detecting spam messages so that it results in reduction of the possibility of threats from the spam messages thereby a fair SMS services are provided by the mobile service providers.

ROAD ACCIDENT ANALYSIS USING MACHINE LEARNING

16X41F0028 | CH.SAIKUMAR

There are many inventories in automobile industries to design and build safety measures for automobiles, but traffic accidents are unavoidable. There is huge number of accidents prevailing in all urban and rural areas. Patterns involved with different circumstances can be detected by developing an accurate prediction models which will be capable of automatic separation of various accidental scenarios. These clusters will be useful to prevent accidents and develop safety measures. We believe to acquire maximum possibilities of accident reduction using low budget resources by using some scientific measures.

Researches are focusing on developing software with the help of machine learning algorithms which can help users to alert about road accidents while travelling. The main objective of this project is predicting the road accidents using machine learning algorithms. Comparative study of the various performances of machine learning algorithms will be done through graphical representation of the results.

The objective and scope of this project is to give information about accidents that occur in different areas, for different ages of people, during light conditions, the most dangerous time to drive. The severity of damage occurring during a traffic accident is replicated using the performance of various machine learning paradigms, such as neural networks trained using hybrid learning methods, support vector machines, decision trees, and concurrent mixed models involving decision trees and neural networks. The experimental results show that the hybrid decision tree neural network method is better than the single method in machine learning paradigms.

DIABETIC RETINOPATHY DETECTION

16X41F0005 | D. VEERA LAKSHMI

Diabetic Retinopathy (DR) is human eye disease among people with diabetics which cause damage to retina of eye and may eventually lead to complete blindness.

Effective treatments for DR are available though it requires early diagnosis and the continuous monitoring of diabetic patients.

Also many physical tests like visual acuity test, pupil dilation, and optical coherence tomography can be used to detect diabetic retinopathy.

Convolutional neural networks are used to find patterns in an image. You do that by convoluting over an image and looking for patterns. In the first few layers of CNNs the network can identify lines and corners, but we can then pass these patterns down through our neural net and start recognising more complex features as we get deeper. This property makes CNNs really good at identifying objects in images.

FEATURE BASED OPINION MINING ON STUDENT FEEDBACK

16X41F0003 | B.SHARMILA

Opinion mining technique for classifying the student's feedback obtained during evaluation survey that is conducted every semester to know the feedback of students with respect to various features of teaching and learning such as teaching, assessments, etc.

Student's feedback can help in lecturers understand their students learning behaviour and improve teaching. Student feedback have the features like the grades, enrolment data. Student feedback can be collected using mobile phones and social media, progression rates as well as unstructured data like student opinion expressed through surveys, web blogs, twitter, face book etc...

The extracted and preprocessed datasets can be subjected to various Machine Learning algorithms such as Support Vector Machine (SVM), Naive Bayes(NB),K Nearest Neighbour(KNN) and Neural Networks(NN). In this proposed system, sentiment classification, machine learning methods have been used to classify each question as positive or negative.

WEBPAGE CLASSIFICATION FOR SAFER BROWSING USING MACHINE LEARNING

16FE1F0033 | SHAIK MASTAN VALI

Now-a-days, Internet fraud becomes a large threat, even though users were attracted and reveals the confidential data to take the benefits that offered by the Hackers through webpage browser and applications in mobile devices. Because of its ubiquitous nature, it is also an extremely the main target of attackers.

Attackers mostly target the web browser to either hi-jack or to snoop on the web traffic from it, or exploit it to access the device itself, and the files saved on it. This was happened because of too much dependency on internet for personal finance, business, investments etc.... So, it is necessary to identify fake websites for end user's security. Once the user visits the website, attacker can fetch the confidential information of the user.

Phishing URL detection can be done via proactive or reactive means. On the reactive end, we find services, such as Google Safe Browsing AP13. This type of services exposes a black list of malicious URLs to be queried. Blacklists are constructed by using different techniques, including manual reporting, honey pots, or by crawling the web in search of known phishing characteristics. For example, browsers make use of blacklists to block access upon reaching the URLs contained in them. This implies that web users remain at risk until the URL is submitted and the blacklist is updated. What is more, since most of phishing sites are active for less than a day, their mission is complete by the time they are added to the blacklist.

SMART CAREER GUIDANCE AND RECOMMENDATION SYSTEM

16X41F0018 | MAJETY SATYA AISHWARYA

Technology is changing its wings day by day and expanding in all dimensions, it is essential for every individual to upgrade and update to forthcoming technologies. In particular technology evolution plays a vital role in the software industry and it is essential for every CSE/IT graduate to acquaint with the skills and certifications of forthcoming technologies. Merely attaining a certification may not suit the graduates to attain a skill. Skill is the thing that comes from a passion and Zeal. In order to understand the passion among the graduates to choose the right course and to gain right Skill Set, it is necessary to predict the mindset of the student and this can be demonstrated with the concept of machine learning. Based on their inherent capabilities, interests and knowledge pre requisites if we train the students in the respective areas they can be excel well in those areas and their efficiency can be maximum.

Our Smart Career Guidance Recommendation System is a system developed for recommending skilling courses and certification courses in the CSE/IT domain.

A substantial amount of literature focuses on predicting student performance in solving problems or completing courses. Many Machine learning techniques, such as decision trees artificial neural networks, matrix factorization, collaborative filters and probabilistic graphical models, have been applied to develop student performance prediction algorithms. In this paper we identify and apply the suitable algorithms for Student specific skill oriented course recommendation system in the CSE/IT domain. We present the dataset built using the questionnaire and skill tests to extract the information regarding their interests, abilities.

KEYWORDS: Machine Learning, Course Recommendation System, Skill Prediction.