## DEPARTMENT OF **INFORMATION** TECHNOLOGY



Towards providing world class education in Information Technology with societal focus.

### **MISSION**

M1. Empowering students with state of art knowledge and skills to meet global challenges

To carry out high quality research M2. leading to the creation and commercialization of intellectual property

To build and provide a quality and M3. comprensive technology infrastructure and committed to empower students to contribute technological and social development and progress of society.

### VISION

### [YEAR 2020-21]

## TECHNICAL **E-MAGAZINE**

- 1. Transfer Learning Approach for classification using chest x-rays images
- Multiperceptron Backpropagation
- Neural Network
- Document collection
- 5. A Deep Learning Approach for Driver Activity **Recognition in Intelligent Vehicles**
- Ethereum
- 7. Evaluating summarization, text identification
- 8. Music Generation Using Deep Learning
- Maps from Satellite Images)
- 11. Travel Recommender System
- 12. Brain tumor detection using CNN
- Questions
- - Mel-Spectograms and WaveNet

#### **INDEX**

multi-disease 2. Neural Network Classification of Blood Cell Images using 3. Robust Lane Detection for Self-Driving Cars using Deep 4. Segmentation-Free Keyword Spotting in Historical

6. A Fund-raising platform using Smart contract and

and comparing various for ways classification, titling, topic

9. Image to Image Translation (Generate Standard Layer of 10. Deep Learning approaches for Bearing Fault Diagnosis 13. Teach machine to read and comprehend text and Answer

14. Sarcasm Detection & Classification using Explainable AI 15. Neural Approach For Text To Speech Synthesis Using

Title: Transfer **Learning Approach for** multi-disease classification using chest x-rays images

Guide: Dr Sangeeta Jadhav



Vishal Savade, Shubham Wagh, Sunit Jena, Vivek Ray

#### ABSTRACT

Thorax diseases cause most of the deaths all over the world. They are caused by fungi, bacteria, and viruses. Radiologists find it difficult to identify the disease just by looking at X-ray images as a patient may have multiple diseases which may be overlapped over other diseases. The main aim of this study helps radiologists to detect the disease with the probability of other diseases.

We proposed an architecture in which a deep learning framework is used for the detection of thorax disease using the concept of transfer learning which would reduce the vast time and model complexity. National Institute of Health Chest X-Ray dataset is used for image preprocessing which contains over 1 lakh images of around 30000 unique patients with 14 different types of thorax diseases, downscaled to 224\*224\*3 which are further augmented and fed to different neural network models pre-trained on ImageNet Dataset. We prepared three different models DenseNet121, MobileNet, and InceptionV3, and we analyzed the performance. We used two ensemble models - voting classifier and stacking

Stacking Classifier or Stacked Generalization is an ensemble machine learning algorithm. It uses a meta-learning algorithm to learn how to best combine the predictions from two or more base machine learning algorithms. It involves combining the predictions from multiple machine learning models on the same dataset, like bagging and boosting. By comparing both the ensemble model with each other and with pre-trained models we found out that with these ensemble models, we outperforms already present state-ofthe-art models

classifier, for combining the output from all pretrained models. A Voting Classifier is a machine learning model that trains on an ensemble of numerous models. The outputs of pre-trained neural networks are combined into a prediction vector, and majority voting is used to come to a final prediction.

Title: Neural Network **Classification of Blood Cell Images using** Multiperceptron **Backpropagation** 

### Guide : Dr Rahul Desai

2

#### Prasad, Pritika

## Benjamin, Rohit Kumar,

classifier, for combining the output from all pretrained models. A Voting Classifier is a machine learning model that trains on an ensemble of numerous models. The outputs of pre-trained neural networks are combined into a prediction vector, and majority voting is used to come to a final prediction.

Stacking Classifier or Stacked Generalization is an ensemble machine learning algorithm. It uses a meta-learning algorithm to learn how to best combine the predictions from two or more base machine learning algorithms. It involves combining the predictions from multiple machine learning models on the same dataset, like bagging and boosting. By comparing both the ensemble model with each other and with pre-trained models we found out that with these ensemble models, we outperforms already present state-ofthe-art models

#### ABSTRACT

Thorax diseases cause most of the deaths all over the world. They are caused by fungi, bacteria, and viruses. Radiologists find it difficult to identify the disease just by looking at X-ray images as a patient may have multiple diseases which may be overlapped over other diseases. The main aim of this study helps radiologists to detect the disease with the probability of other diseases.

We proposed an architecture in which a deep learning framework is used for the detection of thorax disease using the concept of transfer learning which would reduce the vast time and model complexity. National Institute of Health Chest X-Ray dataset is used for image preprocessing which contains over 1 lakh images of around 30000 unique patients with 14 different types of thorax diseases, downscaled to 224\*224\*3 which are further augmented and fed to different neural network models pre-trained on ImageNet Dataset. We prepared three different models DenseNet121, MobileNet, and InceptionV3, and we analyzed the performance. We used two ensemble models - voting classifier and stacking

### Rakesh

Title: Robust Lane **Detection for Self-Driving Cars using Deep Neural Network** 

### Guide : Dr Ashwini Sapkal



3

### Prashant Singh, Dishant Dilip Pawar, Arti

#### **ABSTRACT**

According to WHO 1.35 million people every year are cut short in road accidents, most of them caused due to human misconduct and ignorance. To improve safety over the roads, road perception and lane detection plays crucial part in avoiding accidents.

Lane Detection is a constitution for various Advance Driver Assisting System (ADAS) like Lane Keeping Assisting System (LKAS) and Lane Departure Warning System (LDWS). It also enables fully assistive and autonomous navigation in self-driving vehicles. Therefore, it has been effectual field of research for the past few decades, but various milestones are yet to be achieved.

Finally, we compare various researchers

approaches with their downfall and limitations to predict accurate lanes under different scenarios, we identify these gaps and suggest an approach that may bridge them.

The problem has encountered various challenging scenarios due to the limitation of resources and technologies in past. In this review, we survey the computer vision-based approaches and the techniques algorithmic that have revolutionized the lane detection problem.

## Title: Segmentation-**Free Keyword Spotting** in Historical Document collection

### Guide : Dr Ashwini Sapkal

4

### Sachin Yadav, Pradeep Kumar, Chhavi Malik,

### **Shashank Sharma**

#### **ABSTRACT**

Keyword spotting is the method of estimating whether the text query occurs in the document or not. In this work, We are presenting an efficient segmentationfree word spotting method, applied in the context of historical document collections, follows the query-by-example that paradigm. In our approach, we are using an autoencoder network for image denoising and image binarisation. For the binarized image, we are creating patches using a patch-based framework followed by a siamese network. A siamese network employs two identical convolutional networks to rank similarity between two input word images. Once the network is trained, it can then be used to spot not just

words with varying writing styles and backgrounds but also to spot out of vocabulary words that are not in the training set. The method proposed is evaluated on the Bengali Handwritten dataset.

## Title: A Deep Learning **Approach for Driver Activity Recognition in Intelligent Vehicles**

### Guide : Dr Rahul Desai



5

### Pankaj Kumar, Ashish Kumar Bisoi, Yogesh Kumar Sharma, Aman Mishra

#### ABSTRACT

The World Health Organization (WHO) reported 1.25 million deaths yearly due to road traffic accidents worldwide and the number has been continuously increasing over the last few years. Nearly fifth of these accidents are caused by distracted drivers. Existing work of distracted driver detection is concerned with a small set of distractions (mostly, cell phone usage). Unreliable ad-hoc methods are often used.

This project proposes a deep learning approach to detecting multiple distracted driving behaviours. In order to obtain more accurate detection results, a synchronized image recognition system based on two cameras is designed, by which the body movements and face of the driver are monitored respectively. In this paper, we present the first publicly available dataset for driver distraction identification with more distraction postures than existing alternatives. Driver decisions and behaviours are essential factors that can affect the driving safety. To understand the driver behaviours, a driver activities recognition system is designed based on the deep convolutional neural networks (CNN) in this paper. Common driving activities are identified, which are the normal driving like looking in front or the side view mirrors while, texting, talking on the phone, operating the radio, drinking, reaching behind, hair and makeup, talking to a passenger are distracted driver activities.

In this project, we focus on driver distraction and propose a method to detect driver distraction. We detect driver distraction using single Convolutional Neural Network model such as ResNet -50. In this project, ResNet architecture has been used for classification. This paper proposes the use of residual neural networks (ResNet) to perform distracted driver behaviour recognition. ResNets are a variant of CNNs that utilise skip-connections to realise the training of very deep networks.

## Title: A Fund-raising platform using **Smart contract and** Ethereum

### Guide : Prof Geeta Patil



### Deepak Kumar Chauhan, Saurabh kumar, Rohit Chauhan, Vidyasagar Himanshu

#### **ABSTRACT**

6

companies are newly born Start-up companies which struggle for existence. These entities are mostly formed based on brilliant ideas and grow to succeed. These phenomena are mentioned in the literature of organization, management, and entrepreneurship theories. However, a clear picture of these entities is not available. This project try to conceptualize the phenomenon, i.e., " start-up ", and recognize the challenges they might face. This project proposes a Smart Contract approach in Ethereum to build a trust between the Investor and the Owner of a start-up i.e., A fund raising platform. Once the investor gives permission to spend the money on request then the Smart contract will automatically execute the code and based on which the owner will get its money to spend. This way the investor and the owner both feel like in control of the

#### **Department of Information Technology**

process. Here, we build a smart contract and deploy it on Ethereum network. Both the parties will act as an actor in the smart contract and all the transactions will be coded in the smart contract and allow to do transactions in crypto-currency. In this project, we focus on the control which is being transferred to multiple parties. So, no one gets a full control over it, and this will prevent any malpractices, such as, take away all money, frauded, etc. In this project, React JS will be used in frontend to create an ergonomic platform, Metamask wallets are used for the transactions, Infura APIs will be used, and smart contacts are based on solidity, and Node JS will be in the backend part of the project.

## Title: Evaluating and comparing various ways for summarization, text classification, titling, topic identification

### Guide : Prof Geeta Patil



### Ajay Kumar, Chinmay Gurjar, Gaurav Sharma, Harsh Tiwari

#### ABSTRACT

Our project focuses on four major tasks in Natural Language Processing (NLP) which are Summarization, Text Classification, Titling, Topic Identification. Text summarization is compressing the source text into a diminished version conserving its information content and overall meaning. Text summarization methods can be classified into extractive and abstractive summarization. In this project we took three techniques of summarization.

We concluded with our implementation that Abstractive summarization methods are much accurate. Text classification, also known as text categorization, is a classical problem in NLP which aims to assign labels or tags to textual units such as sentences, queries, paragraphs, and documents.

It has a wide range of applications including question answering, spam detection, sentiment analysis, news categorization, and so on. In this project we tried with both Machine learning and Deep learning-based models and we find out that deep learning-based models have surpassed classical machine learning based approaches in various text classification tasks. In this project, we provide a comprehensive review of 3 models for text classification.

Topic models have been successfully used for analysing text documents. Topic modelling solves the following type

of problem: you have a set of text documents (such as emails, survey responses, support tickets, product reviews, etc), and you want to find out the different topics that they cover and group them by those topics. In this project we discuss and implement probabilistic topic model Latent Dirichlet allocation. Automatic titling (i.e. providing titles) is one of key domains of NLP. This project provides an approach allowing the automatic titling of texts (e.g. emails, fora, etc.).

Title means to understand a large article or any text document in two or three words. So that we can get the idea about article without reading it and after making an idea we can decide to read the whole article or not. Many times, only based on title we can decide that we should read the thing in detail or not. Furthermore, the results and conclusion of this project for various task could be a useful computation tool for social and business research. It will help in selecting best methods in the specific task.

# Title: Music Generation Using Deep Learning

Guide : Prof Vaishali Ingale



### Divit Adlakha, Anush Mohan, Krishan Kumar, Mohit Gupta

#### ABSTRACT

8

In today's world, its a myth that you need to be a music expert to generate music. Even a person who likes music can produce good quality of music. We all like to listen music and if it is possible to generate music automatically then it will prove to be a new revolution in the world of music industry. Until very recently, all music generation was done manually by means of analogue signals. In recent years though, music production is done through technology, assisted by humans. The task that has been accomplished in the papers is the construction of generative neural network architectures that can efficiently portray the complex details of harmony and melody without the need for human interruption. The usage of deep learning to create solutions in the art domain has garnered a lot of relevance and automated generation of musical sequences has been an active area.

This project deals with the creation of music using ABC notation relying on an LSTM-based architecture. This work is primarily fixated on

music generation using musical structure (notes and chords) to assist learning. Thus, a dataset of ABC notations of many tunes is taken in order to train the model. RNNs enable the model to remember previous members of the sequence to determine the next member, hence why they fit the use-case here impeccably. The music generated is assessed manually on the premise that it obeys musical grammar. So, it finally provides us a sequence of characters in ABC notation as output based on the training of model. The music generated is not expected to be professional studio grade music but a sequence that can hold its own against discerning ears. It covers every aspect from following musical grammatical rules to being unique.

Title: Image to Image **Translation (Generate Standard Layer of Maps from Satellite Images**)

### Guide : Prof Vaishali Ingale



### Rishabh Singh, Pragati Patwal, Harshit Mishra, Saket Kumar

#### **ABSTRACT**

Image-to-image translation is a class of vision and graphics problems where the goal is to learn the mapping between an input and an output image. It can be applied to a wide range of applications, such as collection style transfer, object transfiguration, season transfer, photo enhancement and satellite image to map translation.

Our Domain that we are working on is "Generating Standard Layer of Maps from Satellite Images". Automatically generating maps from satellite images is an important task. Creating maps is a very expensive and time consuming process, yet it is one of the most important sources of curated data. Maps have commercial value to companies in multiple sectors of the economy: ride-sharing companies, food delivery companies, national security agencies, and many other sectors of the economy.

Different techniques for image to image translations such as conditional adversarial networks and conditional co-variational auto-encoders are used which takes as its input a satellite image at a specified zoom level and resolution and produces the corresponding human-readable map for that location. The dataset "pix2pix dataset" used, comprised of

Conditional adversarial networks are a generalpurpose solution to image-to-image translation problems. These networks not only learn the mapping from input image to output image, but also learn a loss function to train this mapping. This makes it possible to apply the same generic approach to problems that traditionally would require very different loss formulations. The Conditional adversarial network architecture is comprised of a "generator model" for outputting new plausible synthetic images, and a "discriminator model" that classifies images as real (from the dataset) or fake (generated). The discriminator model is updated directly, whereas the generator model is updated via the discriminator model. As such, the two models are trained simultaneously in an adversarial process where the generator seeks to better fool the discriminator and the discriminator seeks to better identify the counterfeit images.

satellite images of New York and their corresponding Google maps pages. Training set consist of 1097 images and have 1000 images for validation.

## Title: Deep Learning approaches for Bearing Fault Diagnosis

### Guide : Prof. G.M. Walunjkar



# 10Jagmohan Singh, Neeraj Singh, Bhupendra,Abhishek Kumar Singh

#### ABSTRACT

As we know, to reduce friction between moving pairs a crucial component that is used is rolling element bearing. In order to keep out power drives protected an efficient bearing fault diagnosis system is essential. Thanks to IOT (internet of things), a massive amount of data is gathered from bearing health monitoring systems. The volume, diversity and velocity of data that is collected from IOT is huge with varying nature. The main problem in existing way of bearing fault diagnosis is that we need to have some sort of knowledge before hand in the field of signal processing and features are manually extracted. This limits the capability of fault bearing diagnosis. Now, to increase the efficiency we use deep learning models for data mining from big data.

It helps in monitoring the bearing health more precisely than before. Deep learning methods have an edge over conventional machine leaning methods as we don't need to rely on domain knowledge and human analysis. The features causing bearing fault are extracted automatically that humans can't detect. That's why deep learning has attracted attention toward itself. Deep learning approaches are data hungry, they need a lot of data for their training purpose. One of the major challenges in detecting bearing fault is it data availability as degradation data is collected over time. Some institutions have successfully collected the bearing fault data which has been a great help in developing the model. Deep learning can be used for pattern discovery and useful predictions

# Title: Brain **Tumor Detection**

## using CNN

Guide : Prof. G.M. Walunjkar



### Shalini Negi, Shashank Ojha, Suraj Kumar 11 Singh, Tripurari Kumar

#### **ABSTRACT**

to RGB values) and an image of 4 x 4 x 1 array of matrix of grayscale image.

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of main categories to do image the recognition, images classifications. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used. CNN image classifications take an input image, process it and classify it under certain categories (Eg., Dog, Cat, Tiger, Lion).

Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC) and apply Softmax function to classify an object with probabilistic values between 0 and 1. The below figure is a complete flow of CNN to process an input image and classifies the objects based on values.

Computers see an input image as an array of pixels and it depends on the image resolution. Based on the image resolution, it will see h x w x d(h = Height, w =Width, d = Dimension ). Eg., An image of 6 x 6 x 3 array of matrix of RGB (3 refers

# Title: Travel Recommendation

## System

Guide : Prof. Y.N. Gholap



## MD Zaid Alam, Prince Kumar, Adarsh Singh

### , Amit Kumar

#### **ABSTRACT**

12

Recommender systems are using in many different domains. We mainly focuses on the applications of tourisms A widely searched of the conferences since many years has been made. We provide a detailed and upto-date survey of this field, considering the different kinds of ways, the and diversity of recommendation algorithms, the functionalities offered by these systems and their use of Artificial Intelligence techniques.

Our survey also provides some guidelines for the tourism recommenders and suggests the most promising areas of work in this field for the upcoming years. For access to easy and accurate the information is the heart of our system, so has become a common

in this era of the Internet information overload phenomenon and as such a serious issue for those searching for appropriate information.

Furthermore, various researches have been carried out on how to get information on tourism website more effective. So smart intelligent tourism management system tries to overcome the gap by noting what a tourist perceives as relevant, in terms of connecting to tourism products in tourism websites. This study focuses mainly on content because it is seen as the major factor associated with an effective and smart website.

## Title: Teach machine to read and comprehend text and **Answer Questions** Guide : Prof Sandeep Samleti



# Ayushi Sharma, Sachin Bhadouria, Vijay Yadav, Aman Tiwari

#### ABSTRACT

13

The Machine Reading Comprehension is an effort to teach computers how to understand a passage, and then to be able to answer questions from it. It is a complex NLP problem with a lot of ongoing research. There are a variety of real world use-cases for this, ranging from assistive educational technology to dynamic information retrieval. Teaching a computer to read and answer general questions pertaining to a document is a challenging yet unsolved problem.

In this project, a novel neural network architecture called the Reasoning Network (ReasoNet) for machine comprehension tasks. ReasoNets make use of multiple turns to effectively exploit and then reason over the relation among queries, documents, and answers. Different from previous approaches using a fixed number of turns during inference, ReasoNets introduce a termination

state to relax this constraint on the reasoning depth. With the use of reinforcement learning, ReasoNets can dynamically determine whether to continue the comprehension process after digesting intermediate results, or to terminate reading when it concludes that existing information is adequate to produce an answer. Reasoned achieve superior performance in machine comprehension datasets, including unstructured CNN and Daily Mail datasets, the Stanford SQuAD dataset, and a structured Graph Reach ability dataset.

In this project, we studied about different techniques that are used in machine reading comprehension. We studied and implement R-net, S-net and LSTM techniques. Work with two different variants of reading comprehension i.e. Passage Question Comprehension Multiple and Choice Question (MCQ) reading Comprehension.

# Title: Sarcasm detection & **Classification with Explainable AI** Guide : Prof Rupali Bagate

### Aman Saini, Kajal Sethi, Karish Tomar, 14 Amarjeet Singh

#### **ABSTRACT**

enhance.

Basically Sarcasm is the form of verbal irony that is used to annoy someone. Mainly, people use it to say the opposite of what's true to make someone feel foolish or bad. In this study, we propose a method of sarcasm detection using explainable AI in which first our model detects whether the sentence is sarcastic or not and then due to which word it becomes sarcastic. The main motivation behind determining the sarcasm is to identify the level of hurt or the true intent behind the sarcastic text. Optimal features are to be selected before data passes to classification task. So data pre-processing makes the data clean so that the performance of the classifier will be

## Title: Neural Approach for Text to Speech Synthesis Using Mel-**Spectograms and** WaveNet Guide : Prof Aparna Joshi

4/24/2021



### Devansh Dixit, A. Dhanancheyan, Divanshu Tiwari, Gaurav Bhatt

#### ABSTRACT

15

The model design consists of three main parts, Encoder, Synthesizer and Vocoder. The general idea is to extract the voice essence using a small audio sample. After that comes the synthesis of mel-spectrogram with the extracted voice essence and new speech content from text. Finally, the melspectrogram is used to generate voice, which is similar to the input audio sample but with entered text content. The main feature of this model is that it uses much smaller audio input for sampling, hence lesser training time as compared to the conventional method. The Encoder is based on the" Generalized End-to-End loss function", in which we feed the audio to be sampled through a multi-layered LSTM network. The keyword segment is extracted and voice essence is encoded in the final layer called d-vector. The d-vector is sent to the synthesizer for mel-spectrogram generation. After Encoder comes Synthesizer, which consists of the" Tacotron 2" network, a neural network architecture for speech synthesis directly from text. The system is composed of a recurrent sequence-to-sequence feature prediction network that maps character embeddings to mel-

scale spectrograms. To validate the design choices studies of key components of the system are present and evaluation of impact of using mel spectrograms as the conditioning input to WaveNet instead of linguistic, duration, and F0 features is done. Finally, comes the Vocoder which is a modified WaveNet model to synthesize time domain waveforms from those spectrograms. On our observation that large sparse networks perform better than small dense networks and this relationship holds for sparsity levels beyond 96Sparse WaveRNN are more efficient. Hence, a generation scheme based on sub scaling that folds long sequences into a batch of shorter sequences and allows one to generate multiple samples at once. The Subscale WaveRNN can produce 16 samples per step without loss of quality. This method also offers an orthogonal method for increasing sampling efficiency. Sub scaling lets us generate multiple samples at once in a batch. Also using this compact acoustic intermediate representation allows for a significant reduction in the size of the WaveNet architecture.

#### SEMESTER I



### **Processes related to project identification**, allotment, continuous monitoring and evaluation

### SEMESTER II

**Review 3:** For checking theimplementation Status

Student prepare final project report. Appear for university exam (project demo) evaluated by external/internal examiner.







© Copyright, Army Institute of Technology, Pune - 411015

### E-Magzine : Project Abstract 2020-21 VOL 1

### Thanks.

Dr(Mrs)Sangeeta D Jadhav	HOD & Professor
Mr D G Auradkar	Associate Professor
Dr Rahul Desai	Associate Professor
Dr(Mrs)Ashwini Sapkal	Associate Professor
Ms Vaishali Ingale	Assistant Professor
Mrs Geeta Patil	Assistant Professor
Mr GM Walunjkar	Assistant Professor
Mr Yuvraj Gholap	Assistant Professor
Ms Neelima Walde	Assistant Professor
Mr Sandeep Samleti	Assistant Professor
Ms Rupali Bagate	Assistant Professor
Mrs Aparna S Joshi	Assistant Professor
Ms. Jyoti Taralkar	Computer Programmer
Mr. Baliram Deshmukh	Computer Programmer
Mr. Prashant Pradip Karande	Lab. Assistant
Mr. Sonavane Rohan Suhas	Lab. Assistant
Ms. Shital Omkar Suvarnakar	Lab. Assistant
Mr. Suryakant Kenjale	Lab. Assistant