

**JYOTI NIVAS COLLEGE AUTONOMOUS  
POST GRADUATE CENTRE**



**DEPARTMENT OF MCA**

**TECH-ON-TOP**

**E-JOURNAL**

**ON**

**RESEARCH PROJECT PROBLEM STATEMENT PART-1**



**ISSUE 4, NOVEMBER 2021**

# **A SECURE APPROACH TO AUDIO STEGANOGRAPHY**

This Project Titled “A SECURE APPROACH TO AUDIO STEGANOGRAPHY” is the software developed for hiding information which uses the technology called as Steganography which is the art of hiding information in ways that prevent its detection.

A message in cipher text may arouse suspicion while an invisible message is not. The conventional way of protecting information was to use a standard symmetric or asymmetric key system in encryption.

Steganography can also be used to place a hidden “trademark” in images, music, and software, a technique referred to as watermarking.

Steganography, if however, used along with cryptography then the message will become quite secure as far as cryptanalytic attack are concerned. Now, if this cipher text is embedded in an image, video, voice, etc., it is even more secure. If an encrypted message is intercepted, the interceptor knows the text is an encrypted message. With Steganography, the interceptor may know the object contains a message.

When performing data hiding on audio, one must exploit the weakness of the Human Auditory System (HAS), while at the same time being aware of the extreme sensitivity of the human auditory system.

First the audio file which behaves as carrier file is selected. Then a message or a text file to be embedded in audio is selected. Then a key file is selected. The key file contains characters whose ASCII values are taken for encryption. Then the file is embedded in audio using low bit encoding mechanism and extraction of embedded message is being done vice-versa.

## **METHODOLOGY**

This project proposes a technique to implement steganography and cryptography together to hide the data into an audio file.

The basic idea is to present a combination of LSB (Least Significant Bit) technique of audio steganography along with encryption using AES and Bit flipping so as to achieve high security, high data rate and robustness.

A secure approach to audio steganography is achieved through the combination of the following techniques:

1. **LSB:** It embeds secret data into the Least Significant Bit (LSB) of the audio file.
2. **Bit Flipping:** This method is used to conceal secret data in the audio file. Here a block consists of 2 samples and thereby flipping one or two LSBs of the samples to hide secret information in it.
3. **AES:** To increase the level of security, secret messages are encrypted first using cryptographic techniques. A method for developing key generation in audio steganography techniques using the AES (Advanced Encryption Standard) algorithm is used.

**M Komal Singh**  
**20MCA20**

# CURRENCY RECOGNITION SYSTEM

## PROBLEM STATEMENT

As there are many currencies in the world, Currency recognition has become huge problem for people, (mainly money exchange people) so our goal is to make it easy to identify the currencies of various countries around the world. To make it easy and efficient for the recognition of the note. so with the help of this techniques, we can say the origin, denomination and currency name of various country currencies.

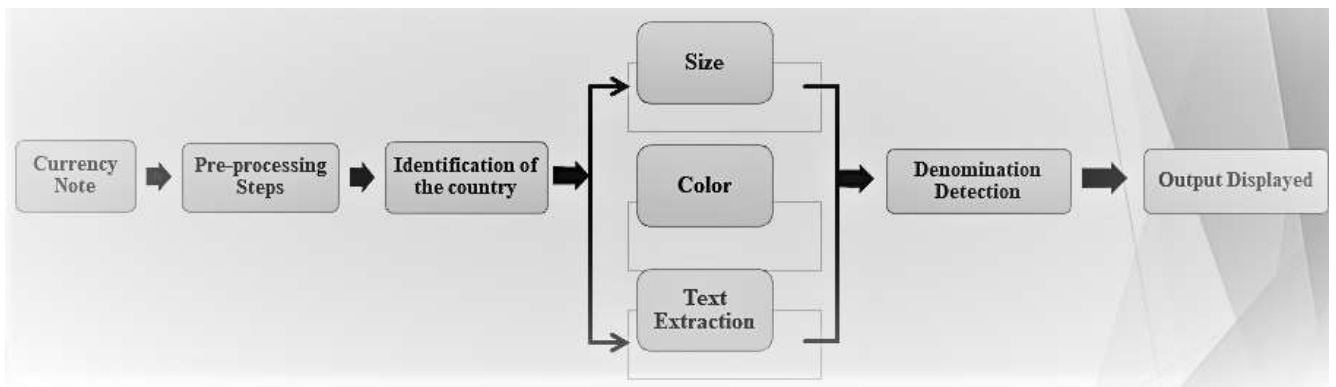
## METHODOLOGY

- K-means clustering
- Template Matching technique

## PRE-PROCESSING STEPS

- Image Acquisition
- Browsing
- Image Preprocessing
- Template Matching

## SURVEY FLOW CHART



**Harshitha M**  
**20MCA**

# **HUMAN POSE ESTIMATION**

## **PROBLEM STATEMENT**

Human Pose Estimation refers to the estimation of the location of the body parts and how they are connected in an image.

Human pose can recognise the pose by looking at the location and positions of different parts of the human body. The same principles is applied in computer vision as problem of Human Pose Estimation is defined as a problem of localization of human joints.

## **METHODOLOGY**

### **1) DEEP LEARNING:**

Human Pose Estimation via Deep Neural Networks DeepPose was the first major paper that applied Deep Learning to Human pose estimation. It achieved SOTA performance and beat existing models. In this approach, pose estimation is formulated as a CNN-based regression problem towards body joints.

### **CURL LOGIC:**

A curl is a tool for transferring data requests to and from a server using PycURL

### **2) OPENPOSE.**

OpenPose is one of the most popular bottom-up approaches for multi-person human pose estimation, partly because of their well-documented GitHub implementation.

**Gopika Sucharitha**  
**20MCA11**

# **IDENTIFICATION OF HATE CONTENTS**

-USING MACHINE LEARNING AND SENTIMENT ANALYSIS

## **PROBLEM STATEMENT**

Building a solution to automatically identify posts on social media platforms that spread hatred of any kind. The solution should be able to automatically learn from any corrections done by users and organically improve over time. Considering posts in English language containing text, images or videos, and ensures that the solution is adaptable to work on other languages in future.

## **METHODOLOGY**

- Machine Learning
- Count Vectorizer Method
- NLP (Natural Language Processing)
- SVM (Support Vector Machine)
- Logical Regression Algorithm

## **ALGORITHMIC STEPS**

Step 1: Importing Libraries Like Pandas, NumPy, sklearn, pipeline

Step 2: Training and testing datasets (Kaggle)

Step 3: installing preprocessor to clean tweets (in order to make our data ready in order to pass into our models)

Step 4: Cleaning the training datasets (tweets) by removing special characters and unwanted spaces.

Step 5: cleaning the test and trained data

Step 6: Import the module from sklearn library and splitting the test and trained data

Step7: training the pipeline mode

Step8: Check the result.

**Anju S M**

**20MCA03**

# IMAGE BLENDING PROCESS

## PROBLEM STATEMENT

This project investigates the problem of fusing multiple images to form a single composite image. Image blending has applications in image editing, panorama stitching, image morphing etc. Human eyes are sensitive to colour and lighting differences within images. Aim of image blending is to provide smooth transitions between image parts which may be obtained from different sources. The simplest image blending method is naive blending which essentially performs cut and paste operation. However this method performs poorly when images to be combined differ in exposure levels, lighting conditions, background colours etc. The transition from one image to another in naïve blending is not smooth and result in undesirable visible seams.

## METHODOLOGY

Generative Adversarial Networks (GANs): are introduced to address the problem of generating realistic images. The main idea of GANs is a zero-sum game between learning a generator and a discriminator. The generator tries to produce more realistic images from random noises, while the discriminator aims to distinguish generated images from the real ones. Although the original method works for creating digital images from MNIST dataset, some generated images are noisy and incomprehensible. GANs must juggle two different kinds of training (generator and discriminator). GAN convergence is hard to identify.

**The Discriminator:** in a GAN is simply a classifier. It tries to distinguish real data from the data created by the generator. It could use any network architecture appropriate to the type of data it's classifying.

- Real data instances, such as real pictures of people. The discriminator uses these instances as positive examples during training.
- Fake data instances created by the generator. The discriminator uses these instances as negative examples during training.

**The Generator:** part of a GAN learns to create fake data by incorporating feedback from the discriminator. It learns to make the discriminator classify its output as real. Generator training requires tighter integration between the generator and the discriminator than discriminator training requires

- Random input
- Generator network, which transforms the random input into a data instance
- Discriminator network, which classifies the generated data
- Discriminator output
- Generator loss, which penalizes the generator for failing to fool the discriminator

**Rachana K V**  
**20mca28**

# **CREDIT CARD FRAUD DETECTION**

## **INTRODUCTION**

Credit cards are used for purchasing goods and services with the help of virtual card and physical card whereas virtual card for online Transactions and physical card for offline Transactions. In a physical card purchase, the cardholder presents the card physically to a merchant for making a payment. In online Payment mode attackers need only less information for doing fraudulent transaction.

## **PROBLEM STATEMENT**

- Credit card fraud detection is to predict fraudulent credit card transactions with the help of machine learning models. As we are moving towards the digital world cybersecurity is becoming a crucial part of our life.
- Credit card frauds are increasing heavily, financial loss is increasing drastically. Every year due to fraud Billions of amounts are lost. Many machine learning algorithms are implemented to detect real world credit card fraud.

## **METHODOLOGY**

### **RANDOM FOREST ALGORITHM**

Random Forest is additionally known for Random Decision Forest (RFA) that will utilized for categorization, Regression with different assignments which is carried out building numerous decision trees.

That Random Forest Algorithm depends onto supervised learning along with significant preferred position for that technique is which that tends to be utilized for categorization and Regression.

Random Forest Algorithm has more good accuracy if contrasted and any remaining existing frameworks with that's the most generally utilized technique.

**Joice Rani J**  
**20MCA16**

# **PLAGIARISM DETECTION**

## **INTRODUCTION**

The plagiarism in student assignments is a widespread and growing problem in the academic process. The traditional manual detection of this kind of plagiarism by human is difficult, not accurate, and time-consuming process.

Many students make (intentionally or unintentionally) some type of cheating and plagiarism in their assignments. Usually it is difficult for teachers to detect plagiarism in student assignments by hand. The detection process becomes easier, faster and more efficient if it's performed automatically (i.e. via a computerized system). The plagiarism detection "is the processes of locating instance of plagiarism within a document wither its text or code". Detection can be either manual or computer-assisted. Manual detection requires effort, and it is impractical in cases where many documents to be compared

## **PROBLEM STATEMENT**

Student assessment is a vital part of the teaching and learning process. The idea of the research project is to propose a plagiarism based system for online paper evaluation. The input answer sheet by the students have to undergo a similarity check with the scheme of evaluation both syntactically and semantically.

## **METHODOLOGY**

The main goal is to create a plagiarism detector. The plagiarism detector will be capable of loading a student's assignment from files and then compute the similarity to determine if students copied each other. Word embedding is used to convert the textual data into an array of numbers because the computer can only understand 0s and 1s.

**Misbah Rafia Khanum**  
**20MCA23**  
**Neha Kouser**  
**20MCA24**



# **HEALTH DETECTION AND DIAGNOSIS FOR COTTON PLANT**

## **PROBLEM STATEMENT**

In cotton, pests and diseases cause vast losses, and these losses are aggravated by unfavorable soil and climatic conditions. The crop is highly prone to insect pests and diseases due to green succulent leaves, long duration crop, Hot and humid weather, and Open flowers etc.

Detecting these diseases with bare eyes is very complex to identify and hence as decreased the crop productivity. At present, the economic loss in a cotton yield ranges from 5% to 15% due to pests and diseases. If no prevention and control measures are taken, the loss may shoot up to 50%. Farmers rely on the heavy use of insecticides to control these pests, but this can cause pesticide residue problems.

## **METHODOLOGY**

A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and are used mainly for image processing, classification, segmentation and also for other auto correlated data. The most common use for CNNs is image classification

In this research we are going to use CNN (Convolution Neural Network) to classify the leaf image and identify the disease. It can automatically detect and recognize diseases based on extracted features at each convolution layer. The image processing technique is used for disease detection. By using CNN technique system can test the image with trained dataset and extract the features.

**Ramya S**

**20MCA30**

# **ABSTRACTION BASED TEXT SUMMERIZATION**

## **PROBLEM STATEMENT**

The primary objective of this problem statement is to prepare a short and accurate summary of news articles on any given topic. An Investigating Officer may sometimes be required to refer to online news articles to obtain further information about a case beyond what is already known through on-ground sources. Due to the proliferation of news websites on the internet, it is not uncommon for a simple search on a topic to return thousands, and even lakhs, of relevant news articles. It would take an Investigating Officer hours and hours of manual effort to go through these news articles, understand them and assimilate key findings.

The concept of automated “text summarization” to reduce the size of document(s) while preserving its meaning is one of the most researched areas among the NLP, or the Natural Language Processing, community. In fact, there are solutions now available in the market that rely on a technique called “extractive summarization” wherein the most important lines and phrases from an article are picked and combined to make a summary. In contrast, “abstractive summarization” promises to produce a really good summary capturing all salient aspects of the source text. The summary is generated using Natural Language Generation techniques and is sometimes even as good as a summary that a human would produce – but in a fraction of time. Build a solution that can quickly prepare a short and concise summary by analysing a large number of online news articles pertaining to terrorism using the “abstractive summarization” technique. The solution should take the desired length of summary from the user as an input.

## **METHODOLOGY/STEPS**

1. Building an Abstract Meaning Representation Graph
2. Generate the graph from the text input such that nodes will contain the information about the POS (part of speech) constrains of the word.
3. Ensuring sentence correctness
4. Getting Abstractive Summary
5. Scoring a Path
6. Scoring is done based on redundancy of the sentences. The redundancy helps in deciding the number of sentences and new score is computed and the duplicate sentences are removed from the fused sentences. The resulting fused sentence and its final score are then added to the original list of candidate summaries. Once all paths have been explored, duplicates are removed.
7. Sentiment Fusion
8. Final Summary
9. Once all the path is scored and the sentences have been fused, we rank the sentences in descending order of their scores. Duplicate sentences are removed for similarity measures using Jaccard similarity.

**Herman K. S**

**20MCA15**

**Boomika S**

**20MCA06**