

# Paddy Crop Disease Prediction- A Transfer Learning Technique



Siddharth Swarup Rautaray, Manjusha Pandey, Mahendra Kumar Gourisaria, Ritesh Sharma, Sujoy Das

**Abstract:** Agriculture is the spinal cord of the human society because it is an essential need of every organism that exists in this planet. Paddy cultivation is very significant so far as humans are concerned, especially in the Asian subcontinent. Since human beings are considered as one of the most intelligent species, it is necessary for us to protect the importance and productivity of agriculture. Since the entry of the IT industry, there has been some improvement in the productivity in the agriculture. It has done a lot of work in the healthcare of the agriculture. Deep learning is a buzzword in the IT sector. This buzzword has helped a lot to improve the productivity of the agriculture field. In the recent past, due to excessive use of human made chemicals and pesticides, the diseases in plants have increased in a higher rate. These diseases in agricultural plants cannot be ignored as it can be dangerous in later stages. Also due to lack of technical knowledge, sometimes it becomes difficult to detect these diseases. So, this paper presents a model for detecting the disease present in the paddy plant. The model uses transfer learning approach which is a paradigm of solving deep learning problems in an efficient manner. This model also finds the probability of the occurrence of disease which can be helpful to take some vital decisions related to plant's health.

**Keywords :** Transfer learning, Convolutional Neural Network, Deep learning, Paddy crop disease.

## I. INTRODUCTION

Development in agriculture is considered as one of the most powerful tools to end extreme indigence and boost the shared prosperity. Agricultural development also aims in feeding a total of 9.7 billion people by 2050. The agriculture sector is very significant for incomes of various classes throughout the world. [1]

But there are major risks in agricultural growth, poverty reduction and food security. Frequent climatic changes and other weather factors may lead to severe reduction in yield of

the crops. Climate change not only declines the growth in agriculture but also snatches the job opportunity of many farmers. Another factor that plays a major role in destruction of agriculture is human error like adding excessive pesticides and insecticide that not only destroys a crop but also harms the biodiversity. Another reason that plays a very nasty role in affecting the agriculture is disease. The disease in plant is as critical as a person suffering from a disease. If not cured in time, could lead in the ultimate demise of the plant.

Paddy crop plays a very major role providing employment to people as farmers in the Asian subcontinent. It not only provides employment, but also helps in eradicating poverty up to some extent. Rice is widely eaten in the Asian sub-continent. It is also considered as staple food in more than 100 countries worldwide. [2] In most of the households it is included with more than one meal per day. It is especially low cost, making itself available to everyone. It is high in calorie and starchy.

But recently there have been some challenges in growing paddy crop. Climate change like global warming, diseases like bacterial, viral and fungal are acting as a road block in the process of cultivating paddy. If not properly examined the factor that the rice plant might be suffering from could let in untimely demise of the rice plant.

The proposed model detects what disease the plant is suffering from. It classifies the disease that the plant might be suffering from into three categories that are Hispa, brown spot and leaf blast. If the leaf is healthy then our model predicts that the leaf is healthy. The model uses Convolutional Neural Network for prediction and classification of the diseases. Convolutional Neural Network is chosen because of its efficiency to work with images. The data set which is chosen contains four classes out of which three are disease class and one is healthy class. The proposed model uses image augmentation for the augmentation of the image and then trains on it to give the desired result.

The proposed model is coded in Python Programming language as it is a best choice for deep learning. The proposed model is run in google's cloud platform that is Google colab.

The organization of the paper is as follows :- Section I is the Introduction, section II is the discussion about the related work, section III is the gives the idea about background study which describes the disease description in a narrowed manner, section IV gives knowledge about our proposed model, section V represents the implementation of our model, Section VI compares the model used in previous implementations to find the diseases that the paddy plant is suffering from, section VII discusses about the result got after the successful implementation of the model and section VIII concludes this paper.

Manuscript received on February 10, 2020.

Revised Manuscript received on February 20, 2020.

Manuscript published on March 30, 2020.

\* Correspondence Author

**Siddharth Swarup Rautaray\***, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India. Email: siddharthfcs@kiit.ac.in.

**Manjusha Pandey**, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India. Email: manjushafcs@kiit.ac.in.

**Mahendra Kumar Gourisaria**, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India. Email: mkgourisaria2010@gmail.com.

**Ritesh Sharma**, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India. Email: sharmaritesh3312@gmail.com.

**Sujoy Das**, School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, India. Email: [dasujoy1999@gmail.com](mailto:dasujoy1999@gmail.com).

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## II. RELATED WORKS

Suraksha et.al [5] proposes a technique which predicts the disease that the paddy crop is suffering by using data mining and image processing techniques. In the paper they have proposed a model for detecting diseases that the paddy plant is suffering from using feature extraction and data mining techniques.

R Rajmohan et.al [6] has proposed a technique which detects the disease that the paddy crop is suffering from by using CNN and SVM classifier. Their model uses feature extraction and SVM classifier for image processing. They have taken 250 images. They have used 50 images for testing and the rest 200 images for training the model. They have also developed a mobile app which clicks the image of the infected plant, zooms it and crops it and then uploads the image and the person receives a notification.

Lipsa Barik [7] has proposed a technique for region identification of Rice disease using image processing. The author has proposed a model which not only identifies the disease that the rice plant is suffering from but also identifies the affected region. The author has used image processing and for classification the author has used machine learning techniques like Naïve Bayes and Support Vector machines. After the prediction is done the severity of the disease is found out and then it is classified into different category.

Anuradha Badage [8] has presented a model for the detection of the disease that the plant is suffering from. The author has used canny edge detection algorithm. The author has used the canny edge detection algorithm to track the edge and get the histogram value to predict the disease that the plant is suffering from. The model also periodically monitors the cultivated field. The model detects the disease in the early stages. Then machine learning is used for training. Then the model takes proper decision and predicts the disease that plant is suffering from.

K.Jagan Mohan et.al [9] presents a model for detecting disease the paddy plant is suffering from. They have used Scale Invariant Feature Transform which extracts the features. The features are then taken and with the help of the features the model detects the disease with the help of SVM and K Nearest Neighbors classifiers.

Dhaygude S. B. et.al [10] presents a model which predicts the disease that the rice plant is suffering from. There are four steps. In the first step RGB transformed image is generated. Then HSI images are generated using the converted RGB. Then the masking of green pixels is done by taking threshold as a parameter followed by segmentation of images and useful feature extraction. At last the texture satisfied is computed from SGDM matrices. After that the evaluation of presence of disease in the leaf is calculated.

Jayanthi, G. et.al [11] proposed a model on analysis of automatic rice disease classification using Image processing techniques. Their paper has presented a detailed study of the different image classification algorithms.

S.Nithya et.al [12] proposes a model using big data. A symptom based recommendation system has been made by them based on the paddy crop disease. The information of diseases are collected from numerous websites and blogs. The information has been analyzed through Hadoop, hive tools and HiveQ. The documents are collected and are represented in vector form using vector Space model and the weight is calculated based on the T-IDF ranking.

Arumugam, A. [13] proposes a model which follows a predictive modeling approach. This model aims to improve the paddy crop production by data mining techniques. Their work aims in providing a predictive modeling approach which will help the farmers to get high yield of paddy crops. They have used machine learning techniques like clustering and Decision trees. They have applied them to the meteorological data.

## III. BACKGROUND STUDY

The objective of this research is to predict what type of disease the rice plant is suffering from. The diseases that the paddy crop is suffering from are described below.

### A. Brown Spot

This disease is the black spots that form on the leaves of the rice plant. This disease can be identified by its symptoms like death of seedlings, death of large areas of the leaf, brown spots or black spots. It comes under the fungal class. It causes both quantity and quality loss. It causes 5% yield loss all over the South and South East Asia. We can make sure the rice plant is not suffering from this disease by providing the crop with correct amount of nutrients and by avoiding water stress. Treating seeds with chemical can also be helpful as it decreases the chance of infection.



Figure 1: Brown Spot infected leaf

### B. Leaf Blast

Leaf blast is considered as the most severe and destructive disease in the rice plants. It can affect various parts of leaves which includes neck and leaves. This disease is very frequent in the regions where there is periodic rainfall, cool temperature and low soil moisture. A Rice plant can have leaf blast at any stage. It can be identified as they appear spots with dark green borders. They are easily confused with brown spots. If grown, it can kill the entire leaf. They are easily confused with brown spots. It can be managed if the field is flooded as often as possible. It can also be managed by splitting nitrogen fertilizer application in two or more treatments. But the excessive use of the fertilizer can increase blast intensity [4] .



Figure 2: Leaf Blast infected leaf

**C. Hispa**

This disease can be identified if the mining of the grubs are clearly seen on the leaves. If the field is severely infested the rice fields appears burnt. If the rice plant is severely infested then the damaged leaves wither off. This disease usually affects the plants in the young stage. This disease can be managed by avoiding over fertilizing the field [3].



**Figure 3: Hispa infected leaf**

**D. Sheath blight**

It is a disease belongs to fungal class. It usually occurs throughout the major paddy growing areas in subtropical and tropical countries. It is found in all areas of rice production, and is decreasing the production of rice especially in intensified production systems. Studies show that it causes a total yield loss of 6% in the tropical Asia [15].



**Figure 4: Sheath Blight infected leaf**

**E. Bacterial Leaf Blight**

Bacterial leaf blight can be found in the leaf of a rice plant and can easily be recognized by observing the yellow and white strips on the leaves. We can know that the plant is suffering from bacterial leaf blight by looking at the youngest leaf which will be pale yellow in color. This disease can be avoided by not using excess nitrogen fertilizer and by plowing stubble and straw into the soil after harvesting the crop. [15]



**Figure 5: Bacterial Leaf Blight infected leaf**

**F. Grassy stunt**

This disease can be identified looking at leaves if they are narrow pale green in color or yellow in color. It belongs to virus class. It can also be recognized if there are irregular dark brown spots on leaves. It can be transmitted by the leaf hoppers. It is a wide spread disease in South, South East Asia, Southern China, Southern Japan and Taiwan [15].



**Figure 6: Grassy stunt infected leaf**

These are the variety of disease that the rice plant might be suffering from. Out of all these diseases we have taken Brown spot, Leaf Blast and Hispa diseases into consideration. Our model can successfully classify these diseases. If the rice plant is not suffering from any disease them it would classify them into healthy class.

**IV. OUR PROPOSED MODEL**

Neural Networks are a series of complex algorithm that finds a primary relationship in a set of data through a process that copies the way the human brain operates. Neural Networks are widely used in planning enterprise, trading, and business analytic, for medical reasons as well as product maintenance.

They have achieved a wide spread adaptation in business application such as fraud detection and it also plays a big role in detecting diseases such as brain tumor, breast cancer as well as diabetes. [5]

Out of different types of neural networks Convolutional neural networks it best suited for working with images. There are a lot of architectures of Convolutional neural networks out of which some commonly used architecture are Alexnet, LeNet-5, VGG 16, Inception, ResNet, and DenseNet. These architectures serve as a rich feature extractor which can be used for image classification, object detection, image segmentation and many other advance tasks.

Out of all these architectures we have used the Transfer Learning approach because it is very convenient to work with a large number of images. Transfer learning is a machine learning approach that focuses on solving a problem by the model's previous experience on different problem. In other words, it is a popular deep learning approach which uses a pre-trained model which is trained over another problem and it solves problems of wide domains like Image Processing, Computer Vision and Natural Language Processing. There are two common approaches to use transfer learning which are Developed Model Approach and Pre-Trained Model Approach.

Developed Model Approach follows four important steps which must be followed to use transfer learning which are selecting source task, developing source model, reusing the model and tuning the model. Pre-trained Model Approach follows three important steps which must be followed in order to use the pre-trained model approach which are selecting the Source model, reusing the model and tuning the model. [14]

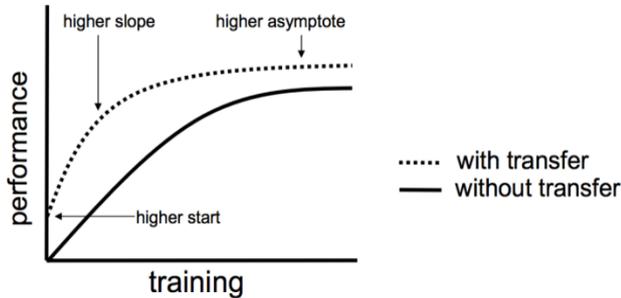


Figure 7: With Transfer v/s without transfer

Transfer learning is commonly used to perform predictive modeling problems that use image data as input. The input may be a photograph or video data as input. Transfer learning is used commonly because it usually gives better performance when compared with other CNN models. There are three main benefits of using Transfer Learning approaches viz. higher start, higher slope and higher asymptote [14].

V. IMPLEMENTATION

For our model we have used the VGG 16 architecture. In VGG 16 the number 16 means 16 layers that have weight.

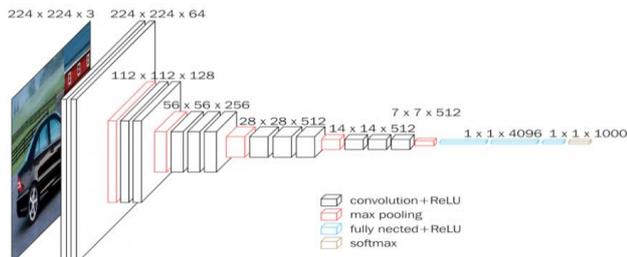


Figure 8: VGG-16 Architecture [17]

A. Image augmentation

For better performance of the neural network, we need a large training dataset which gives good learning experience to the network. Image augmentation techniques are used to virtually increase the size of training data which helps the neural network classifier to perform well. It creates training images artificially by using different ways of combination of multiple processing which are arbitrary rotations, shifts like left shift and right shift, shears and flips.

B. Convolution Step

The word convolution, in the term of mathematics is function which is derived from two provided functions by integration which expressed how the shape of one is modified by the other. In the convolution step a kernel matrix is taken and is multiplied with image matrix to obtain the feature matrix. It contains 2 Convolutional Layer of 64 channel of 3x3 kernel and same padding, 2 Convolutional Layer of 128 channel of 3x3 kernel and same padding, 3 Convolutional Layer of 256 channel of 3x3 kernel and same padding, 6

Convolutional Layer of 512 channel of 3x3 kernel and same padding, totaling to 13 Convolutional Layers. Rectified Linear Unit activation function has been added to each layer so that all the negative values are not passed to the next layer.

C. Pooling Step

It is considered as one of most important step in the Convolutional Neural Network. It is a down sampling strategy that is used in the Convolutional Neural network. The main objective behind applying pooling is to reduce the spatial size of the image. There are 3 kinds of pooling: Max Pooling, Min pooling and average pooling. In our mode we have implemented the max pooling technique.

D. Full Connection Step

A fully connected layer is used to map a matrix to a one-dimensional vector. Pooling layer in our network produces a matrix as an output which works as the input for fully connected layer. A matrix gets converted into a long vector using a flatten() function and linear operations are performed in the dense() layer. Relu() and Softmax are used as Activation function. The information about the dense layers are 1 Dense layer of 4096 units, 1 Dense layer of 4096 units and 1 Dense layer Softmax layer of 4 units. Relu function has been used for both the dense layer of 4096 units to stop forwarding negative values through through the network. Two unit dense layer with softmax activation has been used in the end as there are 4 classes to predict.

E. Process Flow Model

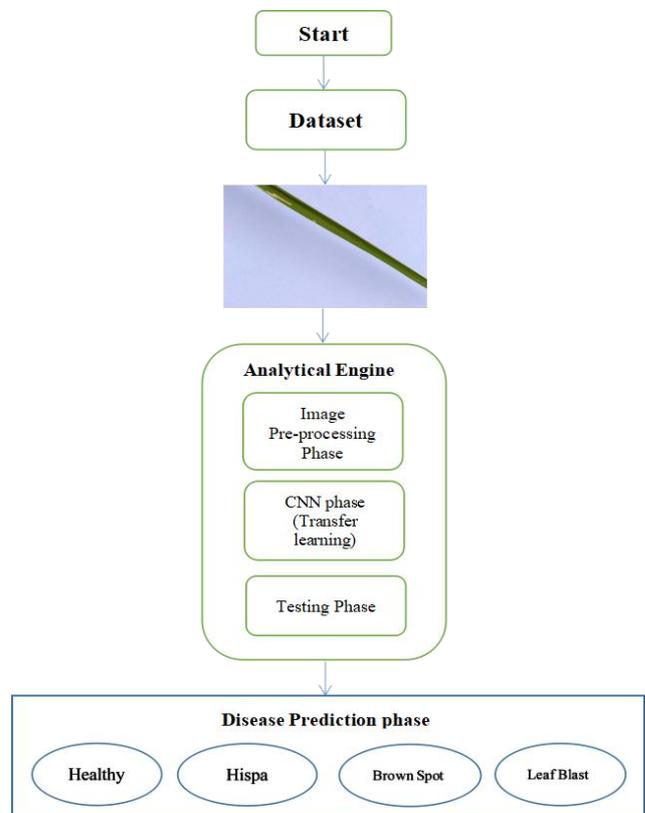


Figure 9: Process Flow Model

**VI. COMPARISON BETWEEN MODELS**

We have tested this model using various architectures of Convolutional Neural Network. We have used one layered CNN, Two layers CNN and have used the 16 layers VGG-16. We have also compared the model with a past model which uses feature extraction to extract features and then KNN and SVM to classify [9]. Our model has achieved highest accuracy that is 94%. The loss in our model is also very less as compared to the other models that we have implemented that are One layer CNN, Two layer CNN. Using transfer learning we have achieved a training accuracy of 92% and testing accuracy of 90%.

**Table-I: Comparison between models**

Model Name	CNN one layer	CNN two layer	KNN and SVM	Transfer Learning
Test Accuracy	72%	75%	70%	91%

**VII. RESULT AND DISCUSSION**

After testing the model with the testing data, the model successfully predicted what diseases the rice plant might be suffering from. We have achieved an accuracy of 90% using the transfer learning method. Transfer learning is an efficient approach to solve complex deep learning problems especially when it comes to image based problems. These set of architectures can be very helpful when we are having a complex problem to solve.

**VIII. CONCLUSION**

Cultivation plays a vital role in the lives of people. It provides means of employment to human beings as well as provides food to humans and animals. As agriculture is important for humans and animals to survive, so it must be protected and preserved. Lack of technical skills and knowledge in the agriculture domain may cause severe issues related to plant’s health. Huge involvement of IT industry in the field of agriculture in the last few years has changed the process of detection and diagnosis of various fatal diseases.

This model uses transfer learning which is a technique for solving complex image processing problems in an efficient and elegant manner. This model uses VGG-16 architecture of transfer learning for prediction of disease in paddy crop. It extracts the feature from the leaves and predicts what type of disease the plant is suffering from. This model strictly follows the principle of classification and detection. The scopes and future of this model is really bright. It can be deployed in a website and an app can be made out of it. The farmer or the person will upload a well clicked image of the plant leaf in the website or the developed app and it will immediately predict what type of disease the plant is suffering from. It will give us the probability of diseases that the plant is suffering from.

**REFERENCES**

1. <https://www.worldbank.org/en/topic/agriculture/overview>
2. <https://www.hsph.harvard.edu/nutritionsource/food-features/rice/>
3. <http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/insects/item/rice-hispa>
4. <http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/blast-leaf-collar>
5. I.S. Suraksha, B. Sushma, R.G. Sushma, K. Susmitha, S.V. Uday Shankar, “Disease Prediction of Paddy Crops using Data Mining and Image Processing Techniques”, International Journal of Advanced

- Research in Electrical, Electronics and Instrumentation Engineering, Vol. 5, Issue. 6, May 2016.
6. R. Rajmohan, M. Pajany, R. Rajesh, D. Raghuraman, U. Prabu, “Crop Disease Identification using Deep Convolutional Neural Networks and SVM Classifier”, International Journal of Pure and Applied Mathematics Vol. 118, No. 15(2018)
7. L. Barik, “Survey on Region Identification of Rice Disease Using Image Processing”, International Journal of Research and Scientific Innovation Vol. 5 Issue 1(2018)
8. A Badage. “Crop Disease Detection using Machine Learning: Indian Agriculture”, International Research Journal of Engineering and Technology Vol. 5 Issue. 9 (2018)
9. K. Jagan Mohan, M Balasubramanian, S Palanivel , “Detection and Recognition of Diseases from Paddy Plat Leaf”, International Journal of Computer Applications Vol. 144 No. 12(2016).
10. S.B Dhaygude, N.P Kumbhar, “Agricultural Plant Leaf Disease Detection Using Image Processing”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2 Issue 1(2013).
11. G Jayanthi.,, K.S Archana., A Saritha, “Analysis of Automatic Rice Disease Classification Using Image Processing Techniques”, International Journal of Engineering and Advanced Technology (IJEAT), Vol-8 Issue-3S pp. 15-20.
12. S Nithya., S Savithri., G Thenmozhi., K Shanmugham, “Symptoms based Paddy Crop Disease Prediction and Recommendation System Using Big Data Analytics”, International Journal of Computer Trends and Technology (IJCTT), 2017.
13. A. Arumugam, “A predictive modeling approach for improving paddy crop productivity using data mining techniques”, Turkish Journal of Electrical Engineering & Computer Sciences, 2017
14. <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>.
15. <https://plantvillage.psu.edu/topics/rice/infos>
16. <https://www.investopedia.com/terms/n/neuralnetwork.asp>
17. <https://neurohive.io/en/popular-networks/vgg16/>

**AUTHORS PROFILE**



**Siddharth Swarup Rautary** presently working as an Associate Professor at School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar Odisha. He did his doctoral degree in Information Technology with MHRD scholarship from Indian Institute of Information Technology (IIITA), Allahabad, U.P. He has an experience of more than 09 years in academia and research. His research interest includes Big Data Analytics; Human Computer Interaction; Data Mining & Warehousing and similar innovative areas. He has published more than 85 research articles in reputed journal, conferences and book chapters. He has delivered 05 invited talks during FDP and in other orientation programs. His research citation index is more than 1600. He has organized number international conferences, workshops and invited talks. Presently he is the faculty coordinator of the data science group at School of Computer Engineering.



**Manjusha Pandey** presently working as an Associate Professor at School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar Odisha. She did his doctoral degree in Information Technology with MHRD scholarship from Indian Institute of Information Technology (IIITA), Allahabad, U.P. She has an experience of more than 09 years in academia and research. Her research interest includes Big Data Analytics; Wireless Sensor Networks, Data Mining & Warehousing and similar innovative areas. She has received the research grant from ICSSR under IMPRESS Scheme for the project title: Intervention of Use of Technology in Government Sponsored ICDS: A Case Study of Khorda District as Project Coordinator. She has published more than 80 research articles in reputed journal, conferences and book chapters. She has delivered 06 invited talks during FDP and in other orientation programs. She has organized number international conferences, workshops and invited talks. Presently she is the faculty coordinator of the data science group at School of Computer Engineering.



**Mahendra Kumar Gourisaria**, is presently working as an Assistant Professor in School of Computer Engineering at KIIT Deemed to be University, Bhubaneswar, Odisha. He has received his master's degree in computer application from Indira Gandhi National Open University, New Delhi and M.Tech. in Computer

Science and Engineering from Biju Patnaik University of Technology, Rourkela. He is pursuing his PhD from KIIT Deemed to be University. He has an experience of more than 15 years in academia and 8 years in research. He has published number of research papers in different International Journals and Conferences of repute. His area of research includes Cloud Computing, Data mining, Soft Computing, Machine Learning and Internet and Web Technologies. He is a member of IAENG, UACEE and life member of ISTE, CSI and ISCA. He can be reached at [mkgourisaria2010@gmail.com](mailto:mkgourisaria2010@gmail.com).



**Ritesh Sharma**, Undergraduate Student, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar is pursuing his Bachelor's of Technology (B.Tech) in Information Technology. He has published a few number of research papers in International Conferences of repute. He is a research enthusiast who is having

more than 2 years of experience in Machine Learning and Deep Learning. He is highly skilled and proficient in various industrial technologies related to Web Technology, Mobile app development, Machine Learning and Deep Learning. His areas of research are machine learning, deep learning and cloud computing. He can be reached out at [sharmaritesh3312@gmail.com](mailto:sharmaritesh3312@gmail.com).



**Sujay Das**, an Undergraduate student, School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar. He is currently pursuing is Bachelor's degree (B.Tech) in Electronics and Computer Science. He is an enthusiastic researcher with an ambition to work in the field of Machine Learning, Deep Learning. He is highly skilled in

leading industrial technologies related to Front-End Web, Machine Learning and Deep Learning. His areas of research are Data Science, Machine Learning and Deep Learning. He is having a 2 years of experience in his domain and has published few number of research papers in reputed conferences. He can be reached at [dasujoy1999@gmail.com](mailto:dasujoy1999@gmail.com)