

# **DAYANANDA SAGAR UNIVERSITY**

**KUDLU GATE, BANGALORE – 560068**



**Bachelor of Technology  
in  
COMPUTER SCIENCE AND ENGINEERING**

## **Special Topic- 1 Report**

### **SENSITIVE DATA CLASSIFICATION USING DEEP LEARNING**

By

<b>MANOJ Y</b>	<b>– ENG20AM0037</b>
<b>MOHAMMED FUZAIL</b>	<b>– ENG20AM0040</b>
<b>MANJU SWAROOP</b>	<b>– ENG20DS0023</b>
<b>SANCHITH S</b>	<b>– ENG21AM3032</b>

**Under the supervision of  
PROF. BINDU MADAVI K.P  
ASSISTANT PROFESSOR**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,  
SCHOOL OF ENGINEERING  
DAYANANDA SAGAR UNIVERSITY,  
(2021-2022)**



**DAYANANDA SAGAR UNIVERSITY**

**School of Engineering  
Department of Computer Science & Engineering**

Kudlu Gate, Bangalore –560068  
Karnataka, India

**CERTIFICATE**

This is to certify that the Special Topic 1 titled “**SENSITIVE DATA CLASSIFICATION USING DEEP LEARNING**” is carried out by **MANOJ Y (ENG20AM0037), MOHAMMED FUZAIL (ENG20AM0040), MANJU SWAROOP (ENG20DS0023), SANCHITH S (ENG21AM3032)** bonafide students of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year **2021-2022**.

**Dr Bindu Madavi K P**

Assistant Professor  
Dept. of CS&E,  
School of Engineering  
Dayananda Sagar University

**Dr Girisha G S**

Chairman, CSE  
School of Engineering  
Dayananda Sagar  
University

**Dr. A Srinivas**

Dean  
School of Engineering  
Dayananda Sagar  
University

Date:

Date:

Date:

**Name of the Examiner  
Examiner**

**Signature of**

1.

2.

# **DECLARATION**

We, **MANOJ Y (ENG20AM0037)**, **MOHAMMED FUZAIL (ENG20AM0040)**, **MANJU SWAROOP (ENG20DS0023)**, **SANCHITH S (ENG21AM3032)** are students of the fourth semester B.Tech in **Computer Science and Engineering**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the Special Topic 1 titled “**SENSITIVE DATA CLASSIFICATION USING DEEP LEARNING**” has been carried out by us and submitted in partial fulfillment for the award of degree in **Bachelor of Technology in Computer Science and Engineering** during the academic year **2021-2022**.

**Students**

**Signature**

**MANOJ Y**

**USN : ENG20AM0037**

**MOHAMMED FUZAIL**

**USN : ENG20AM0040**

**MANJU SWAROOP**

**USN : ENG20DS0023**

**SANCHITH S**

**USN : ENG21AM3032**

**Place : Bangalore**

**Date :**

## ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this Special Topic 1.

First, we take this opportunity to express our sincere gratitude to School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.

We would like to thank **Dr. A Srinivas. Dean, School of Engineering & Technology, Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. Girisha G S, Chairman, Department of Computer Science, and Engineering, Dayananda Sagar University,** for providing the right academic guidance that made our task possible.

We would like to thank our guide **Dr. Bindu Madavi K P, Assistant Professor, Dept. of Computer Science and Engineering, Dayananda Sagar University,** for sparing her valuable time to extend help in every step of our Special Topic 1, which paved the way for smooth progress and the fruitful culmination of the project.

We would like to thank our Special Topic 1 Coordinators, Dr. Savitha Hiremath, Dr. T Kumaresan and all the staff members of Computer Science and Engineering for their support.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Special Topic 1.

# TABLE OF CONTENTS

	Page
LIST OF ABBREVIATIONS .....	i
LIST OF FIGURES .....	ii
LIST OF TABLES .....	iii
ABSTRACT .....	
CHAPTER 1 - INTRODUCTION.....	1
CHAPTER 2 - PROBLEM DEFINITION.....	3
CHAPTER 3 - LITERATURE SURVEY.....	4
CHAPTER 4 - PROJECT DESCRIPTION.....	7
4.1. PROPOSED DESIGN .....	7
4.2. WORKING OF A CNN .....	8
CHAPTER 5 - REQUIREMENTS .....	10
CHAPTER 6 - METHODOLOGY.....	11
CHAPTER 7 - EXPERIMENTATION.....	13
CHAPTER 8 - TESTING AND RESULTS .....	16
CHAPTER 9 - CONCLUSION .....	19
CHAPTER 10 - FUTURE ENHANCEMENTS .....	20
REFERENCES .....	21

## LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ML	Machine Learning
CNN	Convolutional Neural Network
KNN	K Nearest Neighbour
DL	Deep Learning
GUI	Graphical User Interface
GDPR	General Data Protection Regulation
NLP	Natural Language Processing
SVM	Support Vector Machine

## LIST OF FIGURES

Figure 4.1	Overall design
Figure 4.2	Layers of Convolutional Network
Figure 8.1	Accuracy graph
Figure 8.2	Loss graph
Figure 8.3	Final result

# **ABSTRACT**

In the era of Big Data there are a lot of new challenges – understanding, processing, and securing the data, assuring data quality, dealing with data growth and other challenges. One of the challenges is to identify and classify data sets in different systems which must follow the conditions defined by different regulations. The classification of these data sets can be automated using machine learning methods. The aim of the research is to provide machine learning methods for classifying sensitive data. The research is based on analysis and comparison of European Union legislation and scientific literature, which addresses issues of data classification using machine learning methods. Special attention is paid to sensitive data defined by the General Data Protection Regulation (GDPR). The main focus in this research is on supervised learning algorithms, where one of the most effective is Naïve Bayes classifier. In order to achieve good results, there is a need to find a proper training data set. Usage of hybrid methods provides a new way for increasing performance of classifiers



# **Chapter 1 - INTRODUCTION**

Data classification is essential to meet compliance standards and maintain control over sensitive data. Machine learning as a subset of artificial intelligence can help to identify and classify data; therefore, an important aspect is to understand the data regulations and rules that are applied for classifying the data. Sensitive data is a set of special data categories, which must be processed with additional security. These special data categories include religious beliefs, political opinions, racial origin, ethnic origin, as well as membership of professional associations, data about individuals' genetics or biometry and data about sex life and sexual orientation. Machine learning methods are used in many areas because they have the ability to solve complex issues and even make a prediction.

Some of the researchers are working on identifying COVID-19 infection using images of chest X-ray (Sheykhivand et al., 2021). Face recognition (Wang et al., 2019) is also one of the tasks that machine learning can solve. When it comes to classification of text, many of the researchers are trying to identify the hate speech in social networks. In one of such studies Wiedemann, Ruppert & Biemann (2019) developed an architecture of neural network for recognizing the hate speech. Some other related studies (Doostmohammadi, Sameti, & Saffar, 2019) involve subtasks: first task is to identify if the tweet is offensive or not, the other task is to determine whether the tweet is targeted and the last task is to determine to whom it is addressed. The aim of the research is to provide deep learning methods for classifying defined by general data protection regulation.

The system we designed is a deep learning model trained to recognize and categorize PII using Natural Language Processing. The first step of this process was the classification of documents from a dataset that was given to our team mainly consisting of resumes, job postings, and company announcements. Through the determination of document classification, our system then had a set of data that it could learn from. The model is trained to extract relevant features from the text contained in documents. The documents, which come in many formats, are first converted to text, then undergo

preprocessing to extract high value information, and are finally fed into the model to be categorized. We experimented with a number of different models and preprocessing techniques to find the optimal combination.

The GDPR divides data into three different categories:

- Non-Personal Data – Any data that does not directly or indirectly identify an individual.
- Personal Data – Any information that relates to an identified or identifiable person, directly or indirectly. This extends to concrete information regarding identification numbers along with physical, physiological, mental, economic, cultural, or social identity. I.e. if the data can be used to identify an individual based off of the categories above, it is Personal.
- Sensitive Personal Data – Data stored on an individual that identifies said person's racial or ethnic origin, political opinions, beliefs based on religion and/or philosophy, membership of a trade union, and data regarding one's sex-life and health. These categories are tiered in ascending order of Non-Personal Data, Personal Data and Sensitive Personal Data. If both sensitive Personal and Personal data appear in a document then the entire document is considered to be a Sensitive Personal Document.

## Chapter 2 - PROBLEM DEFINITION

**Data sensitivity** concerns information that should be protected from unauthorized access or disclosure due to its sensitive nature. This might include proprietary information about a business that the company wouldn't want its competitors finding out about, or even personally identifiable information about patients or clients.

When our data is uploaded to the internet, there may be chances of leaking personal data, and hence we need to classify our data based on its sensitivity.

**A Machine Learning Model** which can be used by any individual to detect, if their images contain sensitive data or not, thereby protecting the exposure or loss of private and sensitive information.

To protect sensitive data, it must be located then classified according to its level of sensitivity.

The use of such a model will help big companies that work with a large amount of data such that they can prevent any data leaks. A data leak can cost the company thousands of dollars and hence it is very important to safeguard information. If the data is processed with such a model at a larger scale, then we may be able to prevent the damage and hence most importantly lay emphasis on the privacy aspect of data.

With the help of **Deep Learning and Neural Networks** we can train large data sets and classify the data accordingly.

This model can be implemented for various sites such as social media sites. It can prevent sensitive data from being uploaded by warning us about the sensitivity of the information present in a photograph.

## Chapter 3 – LITERATURE SURVEY

[1] In this Deep convolution neural networks are used to identify scaling, translation, and other forms of distortion-invariant images. In order to avoid explicit feature extraction, the convolutional network uses feature detection layer to learn from training data implicitly, and because of the weight sharing mechanism, neurons on the same feature mapping surface have the same weight. The *ya* training network can extract features by  $W$  parallel computation, and its parameters and computational complexity are obviously smaller than those of the traditional neural network. Its layout is closer to the actual biological neural network. Weight sharing can greatly reduce the complexity of the network structure. Especially, the multi-dimensional input vector image WDIN can effectively avoid the complexity of data reconstruction in the process of feature extraction and image classification.

Deep convolution neural network has incomparable advantages in image feature representation and classification. However, many researchers still regard the deep convolutional neural network as a black box feature extraction model. To explore the connection between each layer of the deep convolutional neural network and the visual nervous system of the human brain, and how to make the deep neural network incremental, as human beings do, to compensate for learning, and to increase understanding of the details of the target object, further research is needed.

[2] In this research the main focus is on supervised machines learning algorithms. One of the simplest and most effective is Naive Bayes classifier. It is fast and produces good results. However, when we talk about sensitive data and data regulations, we might need to look at some other algorithms or a combination of algorithms.

An important aspect in machine learning process is training data. Some methods might show better results when are trained using one training data set, but the result might become worse, if using other training data. getting a good training sensitive data is also a challenging task.

In order to achieve even better results, the machine learning methods can be used together with other methods- then they become hybrid methods. This concept, where classifiers are combined, provides a new way for increasing performance of the classifier.

[3] In conclusion, this research is about image classification by using deep learning via framework TensorFlow. It has three objectives that have achieved throughout this research. The objectives are linked directly with conclusions because it can

determine whether all objectives are successfully achieved or not. It can be concluded that all results that have been obtained, showed quite impressive outcomes. The deep neural network (DNN) becomes the main agenda for this research, especially in image classification technology.

DNN technique was studied in more details starting from assembling, training model and to classify images into categories. The roles of epochs in DNN were able to control accuracy and also prevent any problems such as overfitting. Implementation of deep learning by using framework TensorFlow also gave good results as it is able to simulate, train and classified with up to 90% percent of accuracy towards five different types of flowers that have become a trained model. Lastly, Python have been used as the programming language throughout this research since it comes together with framework TensorFlow which leads to designing of the system involved Python from start until ends.

[4] These experiments are carried out from product databases of Amazon, Flipkart, Snapdeal and Paytm. The database currently contains 40000 product catalogues and the classification structure contains 1000 leaf classes. This experiment has been carried out on Intel Core i3 1.80 GHz machine which has 4 GB of RAM. Database server used is MySQL and the application software for implementation of programming code is Anaconda (Spyder 3.6). Approximately 70% accurate results were obtained based on our algorithm and it manages attribute-wise distribution of terms to adapt to the organized way of e-lists.

The best thing is that with the help of normalization, our method without giving weightage to long text is able to give better results. We are in the process of improving the accuracy hence obtained. The algorithm could be made more powerful by including information from more sources. Test information drawn from a more extensive source would likewise give a superior speculation estimate

[5] The model proposed here is for the accurate classification of documents into three different categories, namely, personal, sensitive personal and non-personal documents. The aim of the model is to reduce human efforts of finding the documents containing

an individual's personal information. The organization collecting such information from an individual has the responsibility to protect such documents. A care must be taken by the said organization to avoid leakage of the private information. Failing to which, legal action can be taken on the organization. After countries such as US and EU, India has also strict laws for the security of an individual's private information.

The literature surveyed also proves that Deep Learning model has worked well in document or text classification in numerous applications.

[6] This paper shows the improvement of text classification model by using non-linear features with convolution layer. In starting, input use frequency-based features TF-IDF and semantic features using n-gram features. These features are directly used by machine learning algorithms like KNN, Neural Network and Hybrid network but it does not reduce the non-linearity feature which reduces the accuracy because of lack of information. In proposed approach these features are refined by convolution and reduces the non-linearity which improves the accuracy by 4-5% approximately which indicates that non-linearity of features has important impact on learning.

[7] This paper implements the KNN machine learning classifier, and the results of the data test show that the basic goal is achieved and the classification effect is achieved. The KNN classification algorithm is subjective because a distance scale must be defined. Since the understanding of the distance is not profound, the result of the classification depends entirely on the distance used. Thus, with a set of data, two different classification algorithms will produce two A completely different classification result usually requires experts to evaluate whether the results are valid. Since the recognition of results is often empirical, this limits the use of various distances.

[8] In this paper. we are compiling a new corpus to train the RCNN with, avoiding pitfalls like inconsistent quality, heterogeneous image rights and an inadequate

distribution of image per class. Here we would like to go a dual approach. Together with domain experts, we intend to collate a corpus from the large repository of a major auction house, providing us not only with a selection of artifacts' images but also with texts to be used in multimodal analysis.

On the other hand, this kind of artifacts may exhibit provenance issues (e.g., heterogeneity or lack of provenance). We will thus compensate for such issues by digitizing a major corpus of neoclassical artifacts forming an ensemble and comprising artifacts in multiple modes having evolved in close reference to each other. Therefore, we have entered a partnership with the Dessau-Wörlitz UNESCO world-heritage site, an almost untouched complex of manor houses and their furnishings in early neoclassical style.

Regarding the annotations, we are developing our own semantic annotation and ontology population tool since January 2017. The tool will create annotated corpus. The actual annotation process will be conducted in cooperation with emerging domain experts from the chair of Visual Culture and Art History at the University Passau.

## Chapter 4 – PROJECT DESCRIPTION

The project focuses on identification of data as sensitive or non-sensitive. The approach of CNN is used in order to map the functions and predict the results.

Upon successful training of datasets, we can obtain results for the desired input data. Some data may contain sensitive information and unintentionally may be passed onto the internet by means of social media. This may be harmful because it violates the privacy of an individual. In order to prevent this, we can use the model for safeguarding information.

### 4.1. Proposed design

The model uses Convolutional Neural Network to process the data.

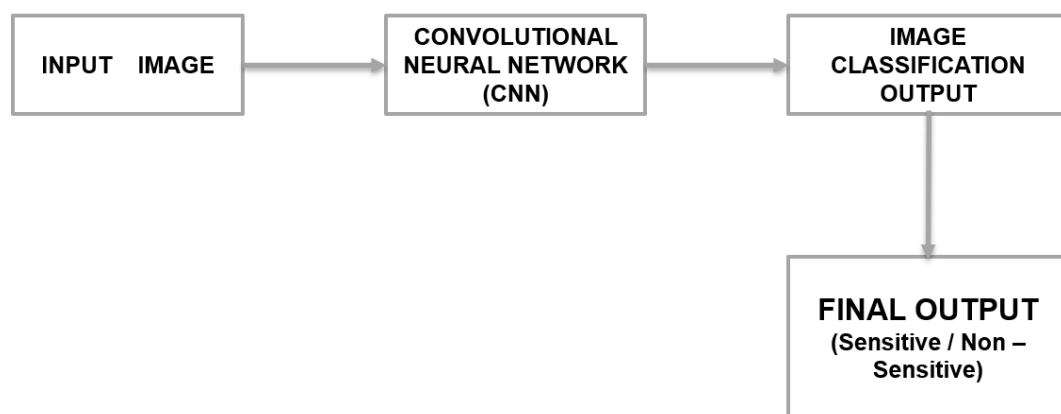


Figure 4.1

The image is fed into the convolutional neural network, which processes the image through its various layers.

The image is then classified according to the criteria mentioned, which in our case is whether it is sensitive or non-sensitive.



## 4.2. Working of a Convolutional Neural Network

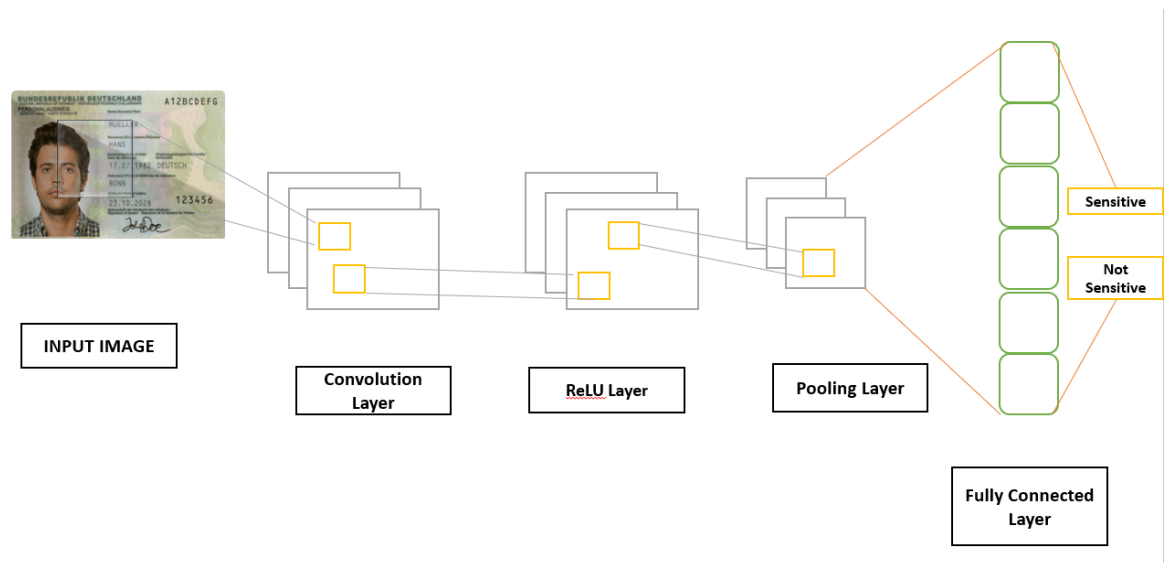


Figure 4.2

From the above image, we can see how exactly the convolutional neural network can be implemented.

CNNs are a class of Deep Neural Networks that can recognize and classify particular features from images and are widely used for analyzing visual images. Their applications range from image and video recognition, image classification, medical image analysis, computer vision and natural language processing.

There are two main parts to a CNN architecture

A convolution tool that separates and identifies the various features of the image for analysis in a process called as Feature Extraction.

A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.

### 1. Convolutional Layer

- This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size  $M \times M$ . By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter ( $M \times M$ ).
- The output is termed as the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

## **2. ReLu Layer**

The purpose of applying the rectifier function is to increase the non-linearity in our images. The reason we want to do that is that images are naturally non-linear.

- When you look at any image, you'll find it contains a lot of non-linear features (e.g., the transition between pixels, the borders, the colours, etc.).
- The rectifier serves to break up the linearity even further in order to make up for the linearity that we might impose an image when we put it through the convolution operation. To see how that actually plays out, we can look at the following picture and see the changes that happen to it as it undergoes the convolution operation followed by rectification.

## **3. Pooling Layer**

- In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations.
- In Max Pooling, the largest element is taken from feature map. Average Pooling calculates the average of the elements in a predefined sized Image section. The total sum of the elements in the predefined section is computed in Sum Pooling. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer.

## **4. Fully Connected Layer**

- The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture.
- In this, the input image from the previous layers are flattened and fed to the FC layer. The flattened vector then undergoes few more FC layers where the mathematical functions operations usually take place. In this stage, the classification process begins to take place.

## Chapter 5 - REQUIREMENTS

Operating System	Windows 11/10/8/7 MacOS Linux
RAM	512MB and above
Processor	Pentium 4 and above
Softwares	Python GitHub GitBash Google Colab Jupyter Notebook
Storage	5GB and above

## Chapter 6 - METHODOLOGY

In this paper, proposed approach depends on deep learning which uses convolution network for features abstraction layer wise and Logistic regression is used on learning part of fully connected layers.

Image classification is one of the hot research directions in computer vision field, and it is also the basic image classification system in other image application fields, which is usually divided into three important parts: image pre-processing, image feature extraction and classifier.

Artificial neural networks constitute a family of machine learning models based on automatically learning data representations. Methods like decision trees learn task specific rules which are used to split the input data into sub-categories at each node. In contrast, neural networks are characterized by not learning any task specific rules. This makes neural networks, while effective, difficult to interpret. The major downside to using neural networks is this lack of explainability. Neural networks consist of an interconnected set of nodes organized into layers, with an input layer whose nodes take on the values of a given input vector, and an output layer which reveals the final computed values. In feed-forward neural networks, considered to be the simplest form of neural network, each node on a given layer is connected to every node on the next layer, and these edges do not form cycles. In this way, the input to any given neuron is a linear combination of the outputs of the neurons in predecessor layer with a set of trained weights. Next, an activation function is applied to the result of this linear combination, which defines the output of a node given an input.

This process may be repeated on successive layers to compute the final output of the network.

The aim of these proposals is to interpret a CNN decision by identifying each time point in the trace that contributes most to a particular classification. However, these techniques are not ones required to understand how the convolutional or classification part selects its feature as they only give a general interpretation of how a network performs. If the network is unable to find the right PoIs, the faulty part (convolutional or classification) cannot be evaluated.

By using a non-linear activation function, neural networks are able to learn to solve nontrivial (non-linear) problems through backpropagation. Backpropagation involves calculation of the gradient of the loss function with respect to the weights in the network in order to iteratively update each of them. Backpropagation is named as such because in networks with multiple layers between the input and the output (deep neural networks) the total error is calculated from the output and then “propagated” backwards through the predecessor layers.

## Chapter 7 - EXPERIMENTATION

The first step of implementation of our code was to link the dataset with the code. In order to classify the data, we need to split the data into individual categories.

The various data categories are:

- X rays
- Policy (insurance, health)
- Passports
- License plates
- Invoices
- Driving licence
- Criminal record
- Cheque
- Biometric data

First, we kept all our data in the same folder, but this method is wrong because we need to have different folders for each category, which is training and validation.

Under training folders, we need to have two sub folders each for sensitive and non-sensitive dataset.

Under validation folders, we need to have two sub folders again. We have implemented the 80% - 20% scheme where 80 percent of the dataset is for training and 20 percent of the dataset is for validation.

In our project, we are using 3000 images out of which 1500 images are for sensitive data and the other 1500 images are for non-sensitive data. Additionally, 500 images are being used for the validation dataset. Out of which, 250 are for sensitive and 250 are for non-sensitive data.

Due to lack of good quality validation images, a low accuracy was seen and hence in order to correct this problem we considered taking good validation dataset because the project depends on how well we can train the algorithm so that we can classify images accordingly. Once this problem was solved, we noticed a better accuracy in the image detection.

There was a problem in the accuracy graph where the plot was incorrectly identified as negative slope. This problem was rectified by changing the parameters of the accuracy function against the value accuracy. Epoch values were obtained for 500 points so that we can ensure that the model is trained with sufficient data.

Finally, we added dataset to GitHub repository using GitBash because the data set was huge. We pushed the data into our repositories to clone the dataset into our code.

Initial code linking was tried using Google Drive cloning but it did not work out as Drive cannot clone datasets correctly into the Google Colab and hence GitHub is the best option for our project.

## Chapter 8 - TESTING AND RESULTS

The model is divided in such a way that it needs to detect if an image is sensitive or non-sensitive. In order to achieve this, we need to have two separate datasets under the categories of sensitive and non-sensitive.

To train the model, we need two sets of data, one being training and the other being validation. The validation phase is followed by the testing phase. The training dataset is larger than the validation dataset.

We can use training dataset to evaluate the performance and progress of your algorithms' training and adjust or optimize it for improved results.

Training data has two main criteria. It should:

- Represent the actual dataset
- Be large enough to generate meaningful predictions

Like we said above, this dataset needs to be new, “unseen” data. This is because your model already “knows” the training data. How it performs on new test data will let you know if it's working accurately or if it requires more training data to perform to your specifications.

Test data provides a final, real-world check of an unseen dataset to confirm that the machine learning algorithm was trained effectively.

In data science, it's typical to see your data split into 80% for training and 20% for testing.

On the other hand, a validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model's hyperparameters.

The validation dataset is different from the test dataset that is also held back from the training of the model, but is instead used to give an unbiased estimate of the skill of the final tuned model when comparing or selecting between final models.



An epoch in machine learning means one complete pass of the training dataset through the algorithm. This epochs number is an important hyperparameter for the algorithm. It specifies the number of epochs or complete passes of the entire training dataset passing through the training or learning process of the algorithm. With each epoch, the dataset's internal model parameters are updated. Hence, a 1 batch epoch is called the batch gradient descent learning algorithm. Normally the batch size of an epoch is 1 or more and is always an integer value in what is epoch number.

Results are calculated based on the accuracy values obtained.

The following graph represents the accuracy versus epochs:

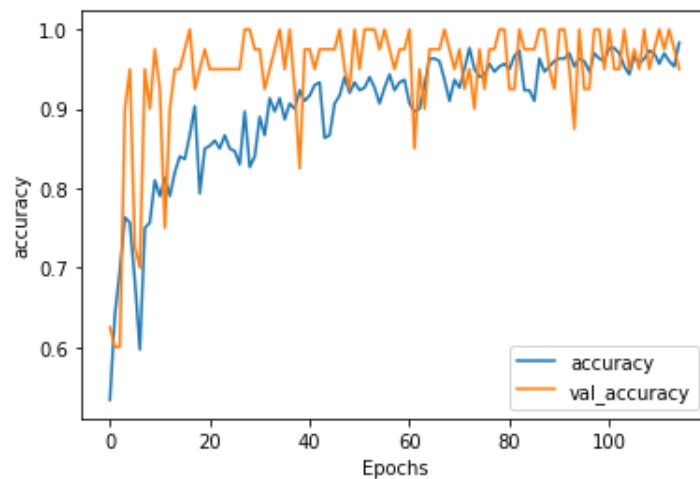


Figure 8.1

The following graph represents the loss versus epochs:

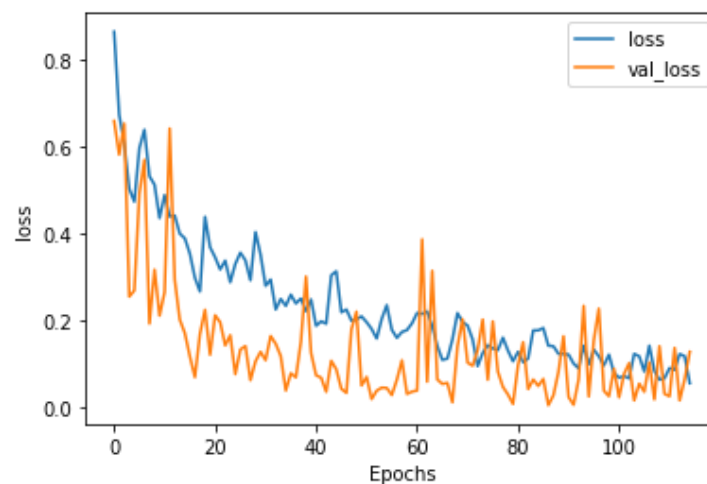



Figure 8.2

Final results can be obtained by choosing the images to be classified:

 Choose Files 2 files

- **aadharcad (122).jpg**(image/jpeg) - 9042 bytes, last modified: 5/23/2022 - 100% done
- **driving\_license (11).jpg**(image/jpeg) - 11837 bytes, last modified: 5/23/2022 - 100% done

Saving aadharcad (122).jpg to aadharcad (122) (2).jpg  
Saving driving\_license (11).jpg to driving\_license (11) (2).jpg  
[1.]  
aadharcad (122).jpg - Sensitive  
[1.]  
driving\_license (11).jpg - Sensitive  
aadharcad (122).jpg: Sensitive



driving\_license (11).jpg: Sensitive



Figure 8.3

## **Chapter 9 -**

## **CONCLUSION**

From the above stated data, we can conclude how a deep learning model can be used to implement this model of sensitive data classification. Out of all the algorithms such as the naïve bayes classifier, random forest algorithm, k-nearest neighbor, CNN, etc., it was found that a Convolutional Neural Network is most suitable for this model of sensitive image classification. CNN is considered because it gives more accuracy as compared to the other algorithms. Also, CNN is helpful in training large datasets in a small amount of time.

Implementation of this project requires a large dataset with good quality validation dataset in order to get maximum accuracy and efficiency. Along with the dataset, additionally there is a need to organize data and separate the sensitive images from the non-sensitive ones so that the model will classify images with a better accuracy. Upon successful training of our dataset, we can upload up to ten images at a time to classify them as sensitive or non-sensitive. The batch size can be increased but that would result in the model taking a longer time to process and classify the images.

Hence, this project is very much useful in the real world as we work with a large amount of data every day. Privacy is the need of the hour and such small steps must be taken to ensure that our information is safe in this world where cyber criminals are always on the lookout to steal personal information and misuse the data.

## **Chapter 10 - FUTURE ENHANCEMENTS**

The project developed may be very useful in the real world in terms of identifying potential sensitive data and classifying that data. This will prevent any further violation of privacy. This project currently takes only a very limited amount of data at a time in order to detect the sensitive data, but with further enhancements we can modify the project so that it accepts a large amount of data at the same time.

This project can be implemented in the form of a website or an application which will be free to use for everyone so that they can use it to find whether a particular image is sensitive or non-sensitive. An application of such type may allow the user to input a large number of images to classify them as sensitive or non-sensitive in a lesser amount of time.

The project can also be improvised by using a better algorithm for CNN with the help of deep learning concepts in order to train larger datasets and get a better accuracy. Efficiency would play a major role in the further development of this project. If a better quality of validation images is used, then we can obtain better accuracy.

Lastly, this project can further be developed as a mobile application which uses the device's camera in order to scan the document and predict whether the document is sensitive or non-sensitive. This would allow for the user to quickly identify sensitive images from anywhere around the world through our mobile phones.

## REFERENCES

- [1] 2019, Ming yuan and Yong Wang. Research on Image Classification model based on Deep Convolution Neural Network
- [2] 2021, Gints rudusans, Gatis. Machine Learning Methods For Classification of Sensitive Data
- [3] 2019, Azlan, Hazirah, Abdul Rahman. A study on Image Classification based on Deep Learning and Tensorflow
- [4] 2018, Shekar, Supriya , Abhilash. Data Classification using Machine Learning Approach
- [5] 2019, Darshan, Reena Lokare. Private Data Classification Using Deep Learning
- [6] 2019, Gitanjali, Kamlesh Lakhwani. Novel Approach of Data Classification using CNN and Logistic Regression
- [7] 2019, Lishan Wang. Research and implementation of Machine Learning Classifier based on KNN
- [8] 2019, Bernhard, Simon Donig, Andre Freitas. Object classification in images of neoclassical artifacts using deep learning