Chandigarh Engineering College Jhanjeri
Mohali-140307
Department of Artificial Intelligence (AI) and Data Sciences

Handwriting Recognition

Project - 44

BACHELOR OF TECHNOLOGY
(Artificial Intelligence and Data Science)

SUBMITTED BY:
Aryan Latotra (2130023)
Badal (2130024)
Baldeep Singh (2130025)
Chandra Prakash Nirala (2130026)

Under the Guidance of
Dr. Rubal Jeet
Associate Professor

Department of Computer Science & Engineering
Chandigarh Engineering College Jhanjeri
Mohali - 140307

Table of Contents

CHAPTER - 1 INTRODUCTION

1.1 Introduction

Machine Learning is a part of Artificial Intelligence (AI). It uses various techniques to provide machines the ability to learn with the help of large amounts of data. While Machine Learning has reduced the burden on programmers to explicitly program the computers, it has touched the areas which have never been explored before. This has led to various technological advances. There are two types of Machine Learning algorithms :

- Supervised Learning Algorithms
- Unsupervised Learning Algorithms
- Reinforcement Learning Algorithms

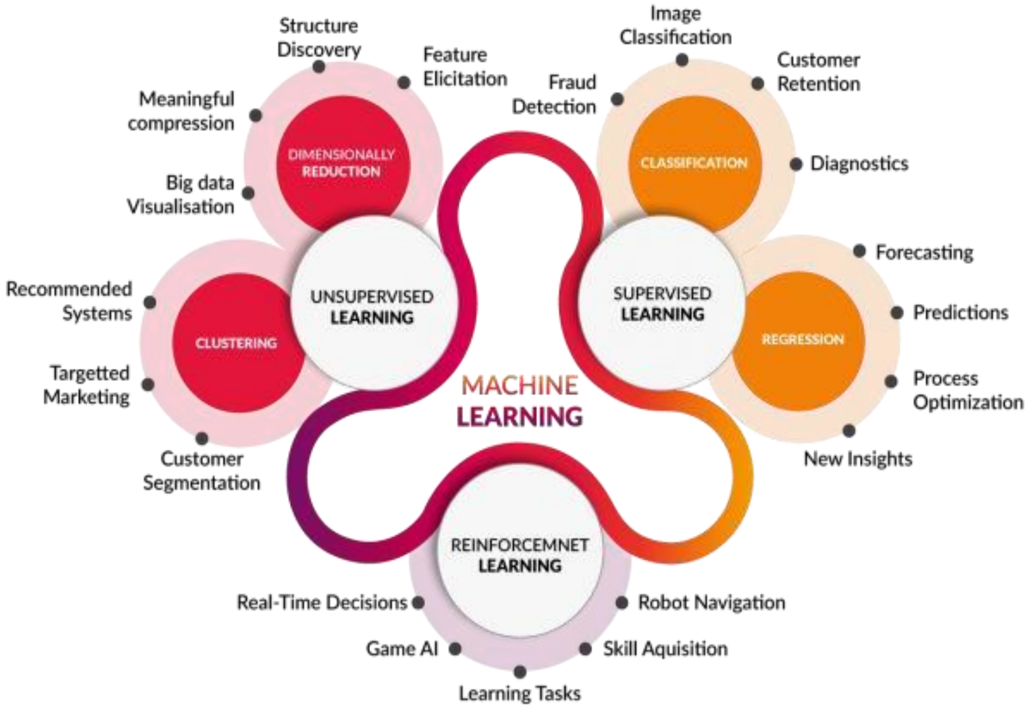


Fig 1.1: Types of ML

In Supervised learning, you train the machine using data which is well "checked." It suggests some data is starting to be marked with the correct answer. It might be stood out from acknowledging which occurs inside seeing a chief or an instructor.

A regulated taking in computation gains from named getting ready data, urges you to foresee results for unforeseen data. Adequately building, scaling, and passing on exact oversaw AI Data science models requires some genuine energy and particular capacity from a gathering of astoundingly skilled data scientists. Furthermore, Data scientists must revamp models to guarantee the encounters given remain substantial until its data changes.

Independent learning is an AI technique, where you don't need to coordinate the model. Taking everything into account, you need to allow the model to manage its own to discover information. It essentially deals with the unlabelled data.

Independent learning estimations license you to perform all the more confounding taking care of tasks diverged from coordinated learning. But, independent learning can be more eccentric differentiated and other typical learning significant learning and stronghold learning methods.

Importance of Machine Learning:

Through cutting edge processing advancements, AI isn't what it resembled previously. It was conceived from design acknowledgment and the hypothesis that PCs can learn without being modified to play out specific assignments. Researchers intrigued by AI needed to check whether PCs could gain from information. The iterative part of AI is significant on the grounds that as models are presented to new information, they can autonomously adjust. They gain from past calculations to deliver dependable, repeatable choices and results.

Simple Programming Model: Machine Learning models can be programmed using Python which is one of the easiest yet the most productive programming languages. Also

Python has one of the biggest developer communities who have provided us with huge libraries which make machine learning algorithms easy to implement.

Cost Effective: Machine Learning with Python guarantees that it doesn't consume a gap in your pocket with regards to overseeing humongous measures of data. This has been an issue with ancestor programming which has been cost restrictive. Numerous organizations have needed to erase and downsize data with the end goal to diminish their expenses.

1.2 Handwriting Recognition:

It's the ability of the computer to interpret and recognise handwritten input. It’s sometimes known as HTR(handwritten text recognition). This could be a scanned handwritten document or a photo of a handwritten note, for instance. The growth and proliferation of touch screens add another way to input handwriting.

The goal of handwriting recognition has been around since the 80s — and has suffered from accuracy issues from the beginning. There are two types of handwriting recognition. First is the older of the two, known as offline handwriting recognition. This is where the handwritten input is scanned or photographed and given to the computer.

The second is online, which is where the writing is input through a stylus/touchscreen. This offers the computer more clues about what’s being written. (For instance, stroke direction and pen weight)

1.3 Problem Statement:

There is an abundance of Handwritten Data which is easily available on the internet. This data can be used for training Machine Learning models which can convert handwritten documents into digital form. This could be advantageous over the following domains:

Healthcare:

Handwriting Recognition can be very beneficial in maintaining the Patient records which are handwritten. This will be a gamechanger for the healthcare industry and make it very easy for patients to access their medical records. It will also decrease the chances of people misreading prescriptions and getting wrong medicines.

Building Databases:

As we all know that paper documents can be destroyed by different means such as floods, fire breakouts and termites. So handwriting recognition will help us overcome all these problems by digitally storing the data over the cloud.

Reducing the cost of storing the Data:

Saving a large amount of data in Physical form can be very costly as it requires huge storage spaces and accessing this data can also be very challenging. Converting the same data to digital format is very useful and cost effective. It will also decrease the use of paper for copying and storing this data will be very beneficial for the environment also.

CHAPTER – 2 BRIEF LITERATURE SURVEY

2.1 The Recognition for Handwritten English Letters

A Review Character affirmation is one of the most captivating and testing research areas in the field of Image taking care of. English character affirmation has been extensively amassed over the most recent 50 years. Nowadays different approaches are in use for character affirmation. File affirmation, progressed library, scrutinizing bank store slips, examining postal addresses, removing information from checks, data area, applications for charge cards, clinical inclusion, credits, charge records, etc are application locales of electronic report getting ready. This paper gives a survey of assessment turn out finished for affirmation of physically composed English letters. In Hand formed substance there is no necessity in the creating style. Interpreted letters are difficult to see as a result of various human handwriting styles, assortment in point, size and condition of letters. Various strategies of physically composed character affirmation are analyzed here close by their introduction.

2.2 Feature Extraction For Handwritten Alphabets Recognition System which is diagonal based Using Neural Network:

A detached interpretation in successive request character affirmation systems using a multilayer feed forward neural association is depicted in the paper. Another strategy, called, corner to corner based component extraction is introduced for eliminating the features of the physically composed letters all together. Fifty instructive records, each containing 26 letter sets created by various people, are used for setting up the neural association and 570 particular translated all together characters are used for testing. The proposed affirmation system performs well indeed, yielding more raised degrees of affirmation 6 accuracy diverged from the structures using the conventional level and vertical procedures for incorporating extraction. This system will be sensible for changing over translated reports into essential substance structure and seeing physically composed names

Using Neural Network for optical character recognition by Image Preprocessing:

Essential errand of this postulation is to make a hypothetical and pragmatic premise of preprocessing of printed text for optical character acknowledgment utilizing forward-feed neural organizations. Show application was made and its boundaries were set by aftereffects of acknowledged investigations.

Handwriting Recognition of Historical Documents:

The majority of current offline handwritten text recognition (HTR) algorithms operate at the line level, converting the text-line picture into a series of feature vectors.These characteristics are supplied into an optical model (for example, a recurrent model).In order to distinguish handwritten characters, a neural network was used. Recent work on document-level text identification and localization and combined line segmentation and identification at the paragraph level has yielded encouraging results. The end outcome However, the finest outcomes in terms of recognition are still to be found. Systems that work at the line level are able to do this.

In this model thirteen stack convolutional layers and three bidirectional layers are used having 256 units in each layer. ReLU is also used to have non- linearity after each layer of CNN. Bidirectional LSTM is used with CTC loss function for making the model end to end trainable. This model was found to be very accurate in predicting on the READ dataset and this model also won second place in ICDAR2017 competition.

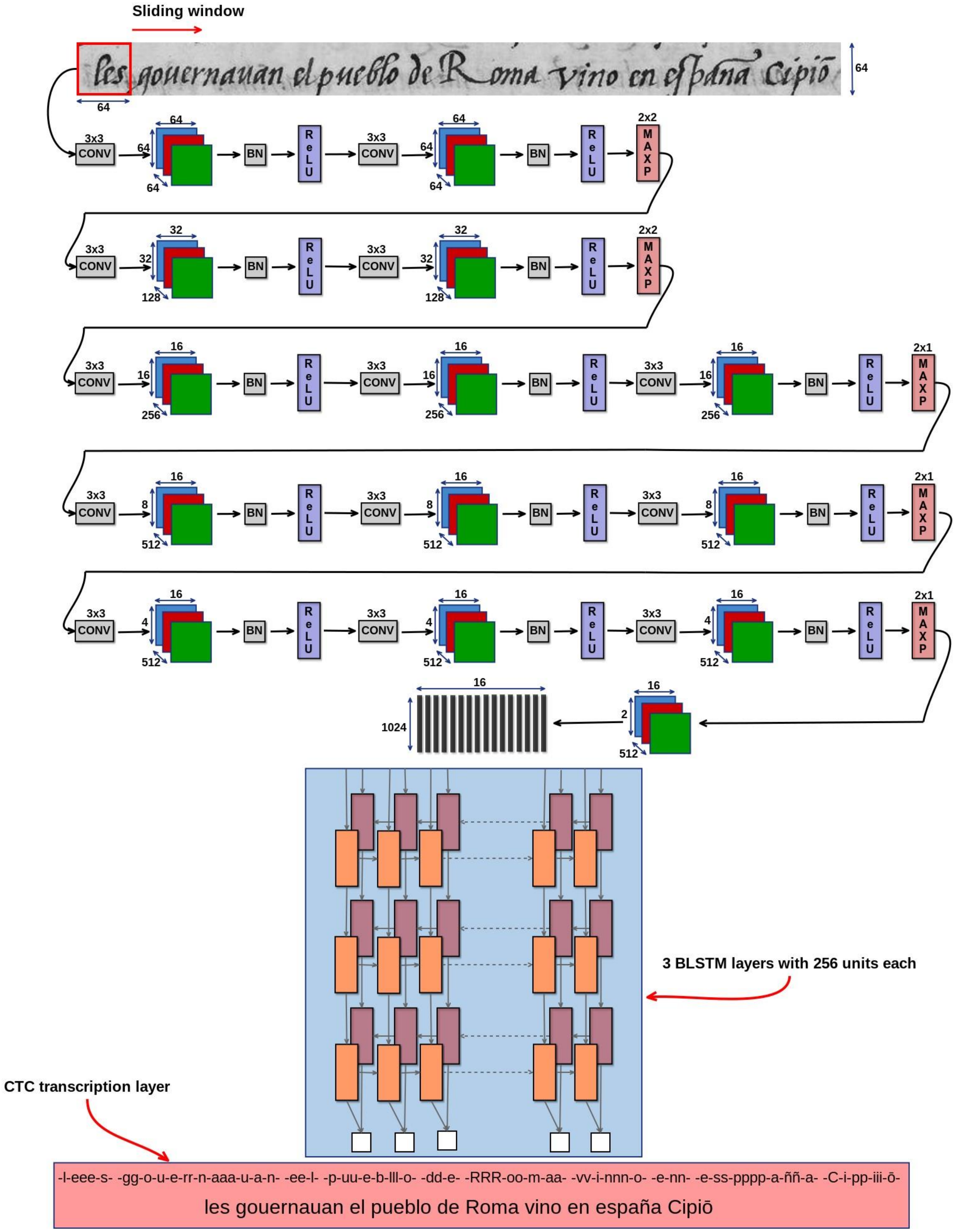

Sliding window
64
64
3x3
CONV
64
64
64
BN
ReLU
2x2
MAXP
32
32
128
16
256
2x1
8
512
4
2
1024
3 BLSTM layers with 256 units each
CTC transcription layer
-l-eee-s- -gg-o-u-e-rr-n-aaa-u-a-n- -ee-l- -p-uu-e-b-lll-o- -dd-e- -RRR-oo-m-aa- -vv-i-nnn-o- -e-nn- -e-ss-pppp-a-ññ-a- -C-i-pp-iii-ō-
les gouernauan el pueblo de Roma vino en españa Cipiō

Fig 2.1: RNN model

CHAPTER – 3 PROBLEM FORMULATION

3.1 Modelling

We are using an incremental model in this project.

What is an Incremental Model?

Incremental Model is a system of programming improvement where prerequisites are broken into various autonomous modules of programming advancement cycle. Steady advancement is done in endeavors from investigation plan, execution, testing/confirmation, support. Each accentuation will encounter the necessities, structure, coding and testing stages. Each resulting appearance of the system adds ability to the last release until the point that all arranged value has been completed.

The structure is placed into age when the essential expansion is passed on. The primary growth is much of the time a middle thing where the crucial necessities are tended to, and worthwhile features are remembered for going with increments. At the point when the middle thing is br0ke somewhere around the cust0mer, there is other plan improvement for the accompanying enlargement.

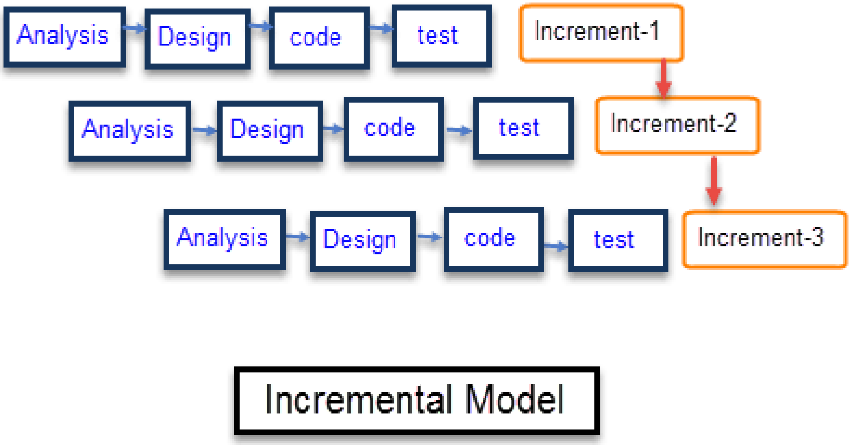


Fig : Pictorial Representation of Incremental Model

3.2 Designing

Volume of information is extending every day that we can manage business trades, sensible data, pictures, chronicles and various others. In this way, we need a system that will be good for isolating the information open and that can normally make reports, viewpoints or overview of data for better use.

There are three phases in designing a model:

Training Data: It is the data on which the machine learning model learns and trains itself. Usually it is large in comparison to test data.

Validation Dataset: Hyper-parameters of a classifier. It is sometimes also called the development set.

Test Data: It is the data on which testing is done and the model is evaluated on the basis of results obtained from this dataset. Usually it is small in comparison to training data.

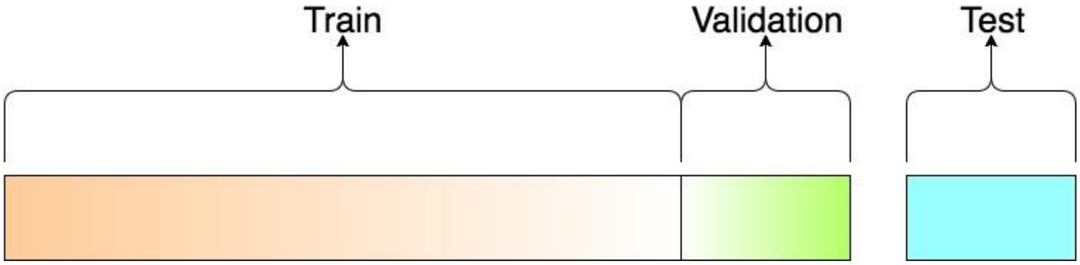


Fig : Three phases of training data

The 7 Steps of approaching a framework in Machine Learning:

Stage-1: Data Collection.

Stage-2: Data Preparation.

Stage-3: Choose a Model.

Stage-4: Train the Model.

Stage-5: Evaluate the Model.

Stage-6: Parameter Tuning.

Stage-7: Make Prediction

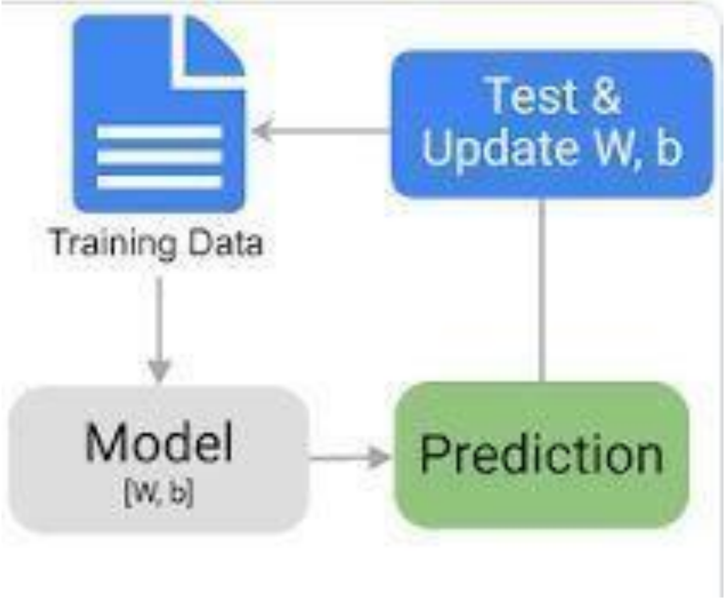


Fig : Frameworks for Approaching the Machine Learning Process

Chapter – 4 Objectives

4.1 Project Objective

The objective of this project is to develop an accurate and efficient handwriting recognition system capable of identifying and converting handwritten text into digital formats. The system aims to achieve the following:

- **Accuracy:** Achieve a character recognition accuracy rate of at least 95% across multiple handwriting styles and languages.

- **Efficiency:** Process handwritten text at a speed of 20 pages per minute or faster, depending on the hardware configuration.

- **Scalability:** Support datasets exceeding 1 million samples, enabling adaptability for large-scale real-world applications.

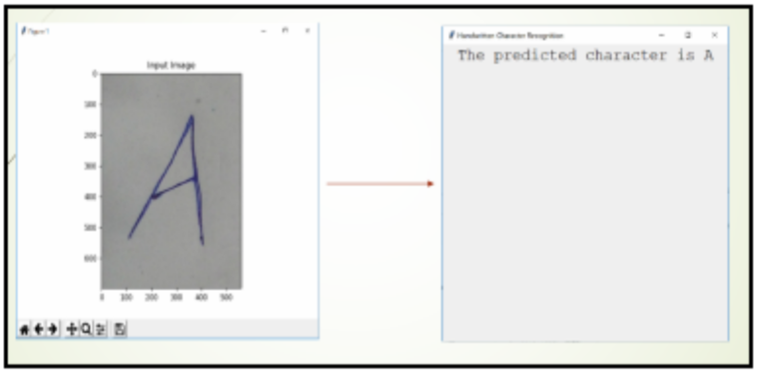
- **Multi-language Support :** Recognize and interpret at least 5 different languages with distinct character sets, including English, Arabic, and Chinese.

- **Platform Independence :** Design the system to be deployable on diverse platforms such as mobile devices, desktop systems, and cloud services.

To achieve these goals, the project will utilize deep learning-based OCR (Optical Character Recognition) techniques, leveraging convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for feature extraction and sequence prediction, respectively.

4.2 Expected Outcomes:

Recognition performance benchmarked against popular datasets like IAM and MNIST, with word error rates (WER) below 10%.
A user-friendly application capable of processing real-world handwriting samples, including mixed-style inputs.



```
The predicted character is A
```

Fig. Character Detection

CHAPTER – 5 METHODOLOGY AND PLANNING OF WORK

5.1 Where to Implement ?

- In machine learning models, we are basically using the CNN (convolutional neural network). A CNN basically consists of an input layer, an output layer, and a hidden layer,

  which can have multiple layers. A convolution operation is performed on these layers using a filter that performs 2D matrix multiplication on the layer and filter.
- We are using the Google Colab for coding, testing, and verification of an existing dataset for character detection.
- The model we used was built with Keras and Sklearn using convolutional neural networks (CNN).
- So the motive of our project is to recognise the handwritten character from an image, paper, document, or any other source. It will save a lot of time for the person, and it will move our society towards digitalization

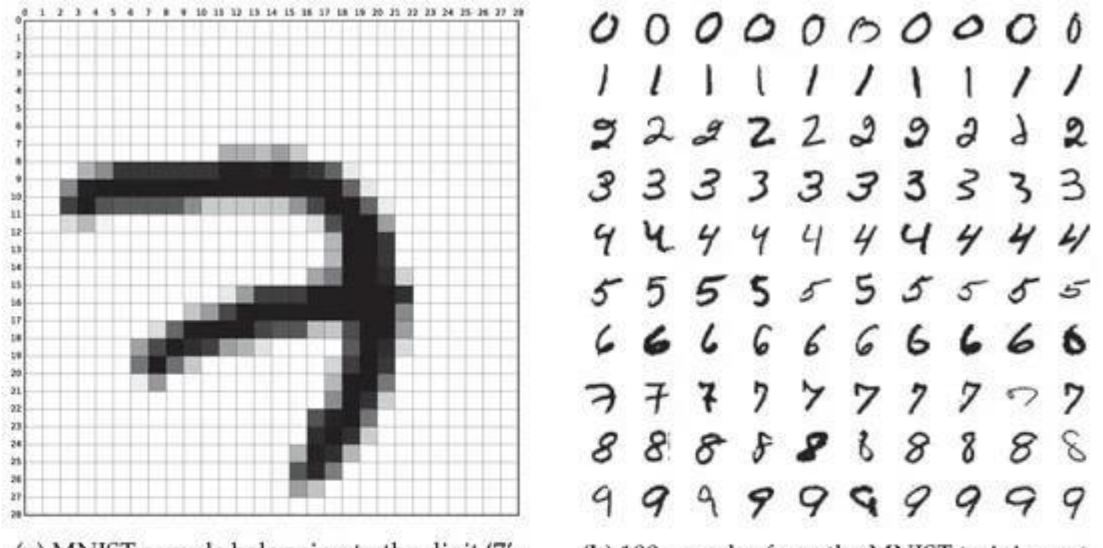
We have used Python for implementing our project.

Why Python?

Python is very easy to understand and is a beginner friendly high level language. Python's simplicity allows us to write reliable systems. Python is more mechanical and models are quickly trained for machine learning.

5.2 Datasets Used:

The MNIST database of physically composed digits, open from this page, has an arrangement set of 60,000 models, and a test set of 10,000 models. It is a subset of a greater set open from NIST. The digits have been size-normalized and centered in a fixed-size picture.


(a) MNIST sample belonging to the digit '7'.

(b) 100 samples from the MNIST training set.

Fig 5.1: MNIST Dataset

Another dataset used in this project is "Handwriting Recognition" dataset. This dataset is a compilation of about 4 Lakh handwritten names. There are 2.06 Lakh first names and 2.07 Lakh surnames in total in this dataset. The data is further divided into a training set of 3.36 Lakh , testing set of 41.83 Thousand and validation set of 41.38 Thousand respectively.

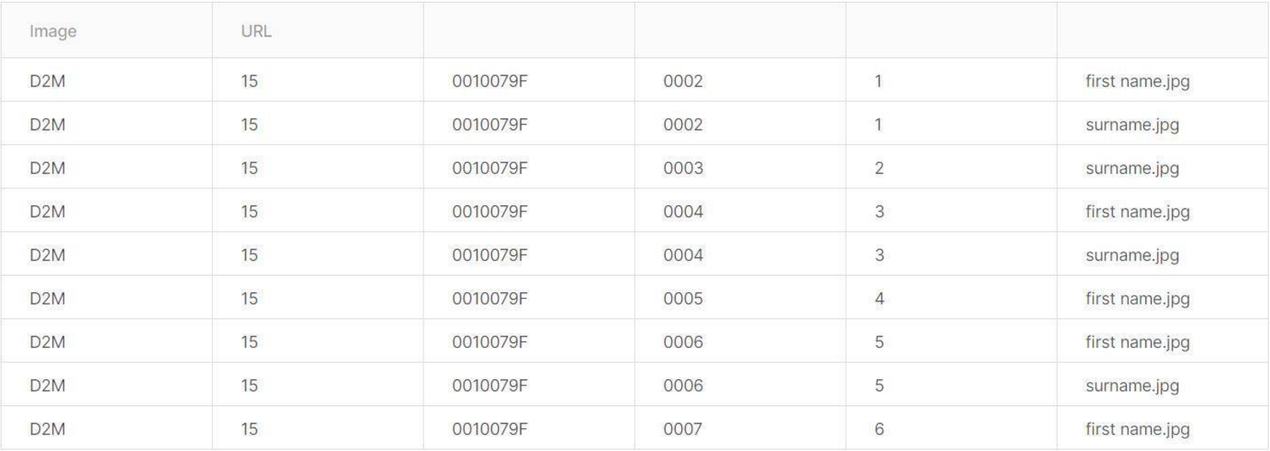


Image	URL				
D2M	15	0010079F	0002	1	first name.jpg
D2M	15	0010079F	0002	1	surname.jpg
D2M	15	0010079F	0003	2	surname.jpg
D2M	15	0010079F	0004	3	first name.jpg
D2M	15	0010079F	0004	3	surname.jpg
D2M	15	0010079F	0005	4	first name.jpg
D2M	15	0010079F	0006	5	first name.jpg
D2M	15	0010079F	0006	5	surname.jpg
D2M	15	0010079F	0007	6	first name.jpg

5.3 Block Diagram

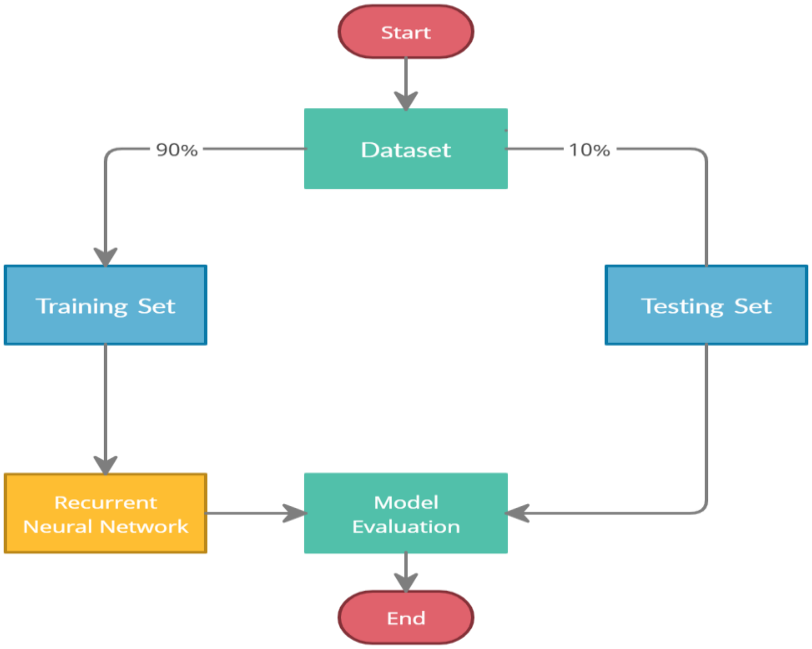


Fig : Block Diagram

The data from the dataset is divided into training and validation sets which are in the ratio of 10:1.

Training set contains 38000 images whereas the testing set contains 3800 images.

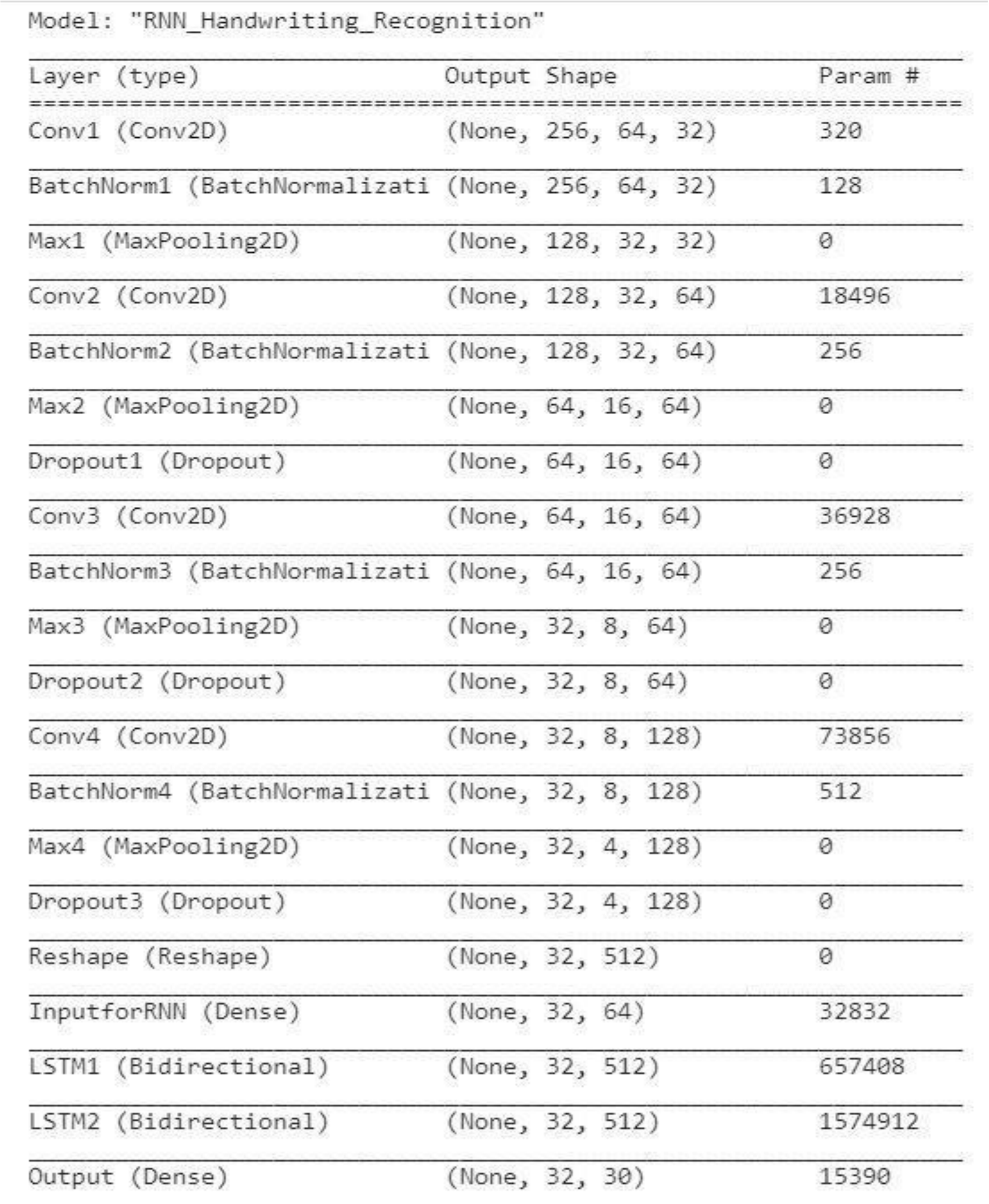
The RNN is trained on the images from the training set and then its accuracy is checked by using the model on the validation set.

5.4 Libraries Used:

We have used following libraries in the implementation code:

- Pandas
- Numpy
- Matplotlib
- Tensor-Flow
- Keras

5.5 RNN Model Implemented:



```
Model: "RNN_Handwriting_Recognition"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
Conv1 (Conv2D)               (None, 256, 64, 32)       320
_________________________________________________________________
BatchNorm1 (BatchNormalizati (None, 256, 64, 32)       128
_________________________________________________________________
Max1 (MaxPooling2D)          (None, 128, 32, 32)       0
_________________________________________________________________
Conv2 (Conv2D)               (None, 128, 32, 64)       18496
_________________________________________________________________
BatchNorm2 (BatchNormalizati (None, 128, 32, 64)       256
_________________________________________________________________
Max2 (MaxPooling2D)          (None, 64, 16, 64)        0
_________________________________________________________________
Dropout1 (Dropout)           (None, 64, 16, 64)        0
_________________________________________________________________
Conv3 (Conv2D)               (None, 64, 16, 64)        36928
_________________________________________________________________
BatchNorm3 (BatchNormalizati (None, 64, 16, 64)        256
_________________________________________________________________
Max3 (MaxPooling2D)          (None, 32, 8, 64)         0
_________________________________________________________________
Dropout2 (Dropout)           (None, 32, 8, 64)         0
_________________________________________________________________
Conv4 (Conv2D)               (None, 32, 8, 128)        73856
_________________________________________________________________
BatchNorm4 (BatchNormalizati (None, 32, 8, 128)        512
_________________________________________________________________
Max4 (MaxPooling2D)          (None, 32, 4, 128)        0
_________________________________________________________________
Dropout3 (Dropout)           (None, 32, 4, 128)        0
_________________________________________________________________
Reshape (Reshape)            (None, 32, 512)           0
_________________________________________________________________
InputforRNN (Dense)          (None, 32, 64)            32832
_________________________________________________________________
LSTM1 (Bidirectional)        (None, 32, 512)           657408
_________________________________________________________________
LSTM2 (Bidirectional)        (None, 32, 512)           1574912
_________________________________________________________________
Output (Dense)               (None, 32, 30)            15390
```

Fig 5.15: RNN Model

In this model there are 4 Convolution layers which are used for feature extraction from the image.
There are 4 Max Pooling layers to get only prominent features.
3 dropout layers are also there so that the model does generalize and overfit the data.
The reshape layer is used to reshape the data from the Covnets and make it suitable to feed it into the RNN.

The RNN layer uses bidirectional LSTM so that words can be predicted accurately.
At the end we have used a dense layer with then outputs to classify the words.

Chapter - 6 Facilities required for Proposed Work

6.1 Software Used:

- Jupyter Notebook
- Google Collab
- Kaggle

6.2 Requirements of System:

Windows 10, Windows Vista, Linux, Ubuntu.

Jupyter Notebook:

The Jupyter Notebook is an open-source web application that licenses you to make and share reports that contain live code, conditions, observations and record text. Uses include: data cleaning and change, numerical reenactment, quantifiable showing, data recognition, AI, and impressively more. Jupyter Notebooks are a fantastic technique to make and rehash on your Python code for data assessment.

Why a Jupyter NoteBook?

They turn into latex reports really easy. They also turn into slideshows really easy. They also let you run blocks of code really easily.

Some other alternatives: PyCharm, RStudio

Google Collab:

It is made by Google’s Research Department. It's used to execute and write Python code online in a browser. It is very useful and efficient for people who are trying to learn Machine Learning.

6.3 Minimum Hardware Requirement:

2GB RAM

Pentium Dual Core 4

Memory Required 20 MB

REFERENCES

[1] Nisha Sharma et al, “Recognition for handwritten English letters: A Review”International

Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 7

[2] J.Pradeep et al,, “Diagonal based feature extraction for handwritten alphabets recognition System using neural network”International Journal of Computer Science and Information Technology (IJCSIT), Vol 3, No 1

[3] Miroslav NOHAJ, Rudolf JAKA, “Image preprocessing for optical character recognition using neural networks”Journal of Pattern Recognition Research

[4] Jehad Ali, Rehanullah Khan, Nasir Ahmad, Imran Maqsood " Random Forests and Decision Trees ".

https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-inpython/

[5] A. Graves, S. Fern´andez, F. Gomez, and J. Schmidhuber,“Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks,” in Proceedings of the 23rd international conference on Machine learning. ACM, 2006, pp. 369–376.