# Scribble to Script: An AI-powered Handwriting Extraction System

~ ROHAN SAI AMPATY

# ABSTRACT

- This work proposes a handwritten text detection system using image pre-processing, a Bi-LSTM character recognition model, and post-processing techniques.
- The system exhibits good performance on diverse test cases with variations in handwriting styles, image quality, and background complexity.
- Future directions include data augmentation, exploring advanced deep learning architectures, and integrating the system into real-world application

SR UNIVERSITY

# INTRODUCTION

- It leverages image pre-processing techniques to prepare the input image, followed by a Bi-directional Long Short-Term Memory (Bi-LSTM) network for character recognition within each word region.
- The system demonstrates promising results on test data with variations in handwriting styles, image quality, and background complexity.
- The document explores potential improvements through data augmentation, advanced deep learning architectures, and integration with real-world applications.

# LITERATURE SURVEY

- Handwritten text recognition (HTR) extracts written text from images. This survey explores techniques like statistical features, structural features, and deep learning (CNNs, RNNs) used for HTR.

- Bi-directional Long Short-Term Memory (Bi-LSTM) networks are powerful for HTR. Bi-LSTMs handle long sequences well, making them ideal for character order in text. Bi-LSTMs process information in both directions, further improving context modeling for characters within a word.

- Pre-processing techniques like grayscale conversion, binarization, and normalization prepare images for the HTR model. These techniques improve text segmentation and feature extraction, leading to better accuracy.

# PROBLEM IDENTIFICATION

.

- Data entry from handwritten forms, surveys, or applications can be automated. Extracted text can be directly fed into databases or spreadsheets, saving time and reducing errors
- Real-time handwritten text capture and conversion on mobile devices can be used for tasks like capturing notes, translating languages on signs, or filling out digital forms with a stylus.
- By deciphering handwritten text in historical documents, researchers and archivists can gain insights into past events and analyze handwritten manuscripts.

# SYSTEMATIC ANALYSIS

- Analyzing how grayscale conversion, binarization, and resizing affect text detection. Comparing accuracy with and without each step.
- Testing the system on handwriting styles, image qualities (clear/blurry/noisy), and backgrounds to see how well it generalizes.
- Adjusting Bi-LSTM hyperparameters (layers, learning rate) and see how it affects performance. Using techniques like grid or random search for optimal settings.

# THE KNOWLEDGE GAP

- Data Bias: HTR models might favor writing styles or languages in their training data, hurting performance on unseen variations. Research is ongoing to tackle bias and ensure fairness.
- Complex Layouts:  HTR systems struggle with complex document layouts or overlapping text. New algorithms are needed to handle these challenges effectively.
- Real-World Integration:  A gap exists in seamlessly integrating HTR models into real-world applications. Research is needed for robust post-processing and integration with other functionalities.
- Explainability: Deep learning models like Bi-LSTMs are powerful but lack transparency. Research is underway to explain their decision-making process for debugging and building trust.

# MOTIVATION AND BACKGROUND

- **Motivation:**
- Automating Information Extraction: Handwritten text detection facilitates automatic extraction of information from images containing text, reducing manual effort and improving efficiency in tasks like document processing and form recognition.
- Preserving Historical Documents: This technology enables digitization and analysis of historical documents with handwritten content, promoting accessibility and knowledge preservation.
- Assistive Technologies: Handwritten text detection can be a valuable tool in assistive technologies for visually impaired users, allowing them to interact with text-based information in images and documents.

# MOTIVATION AND BACKGROUND

- **Background:**
- Handwritten text recognition remains a challenging task due to the inherent variability in writing styles, character formations, and potential ambiguities.
- Deep learning models, particularly recurrent neural networks (RNNs) like LSTMs, have emerged as powerful tools for tackling this challenge by effectively capturing sequential information within handwritten text.
- This project aims to develop a robust handwritten text detection system that leverages image processing techniques and deep learning models to accurately identify and recognize text in various image sources.

# OBJECTIVES

- **Develop a robust handwritten text detection system:** This objective focuses on creating a system that can accurately identify and localize text regions within various image sources.
- **Leverage deep learning for character recognition:** This objective highlights the use of a Bi-LSTM model, a type of deep neural network, to effectively recognize individual characters within the detected text regions.
- **Achieve accurate text detection across diverse scenarios:** This objective emphasizes the system's ability to handle variations in handwriting styles, image quality, and background complexity, ensuring broader applicability.

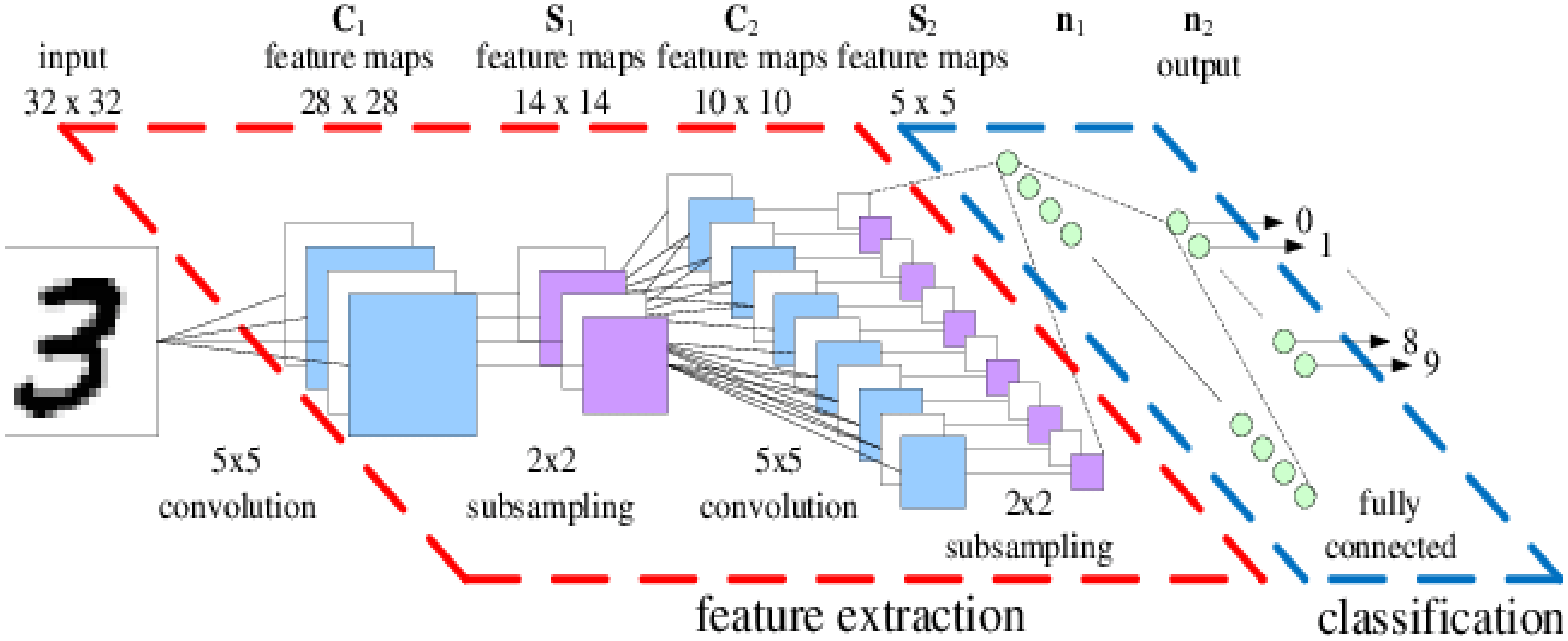
# EXSISTING SYSTEM

- Existing handwritten text detection systems can be broadly categorized into rule-based, statistical learning, and deep learning approaches.
- Rule-based systems are fast and efficient but may struggle with complex layouts or variations in handwriting styles.
- Statistical learning approaches use features like pixel intensity or connected component properties for classification, offering some adaptability but potentially missing complex data relationships.

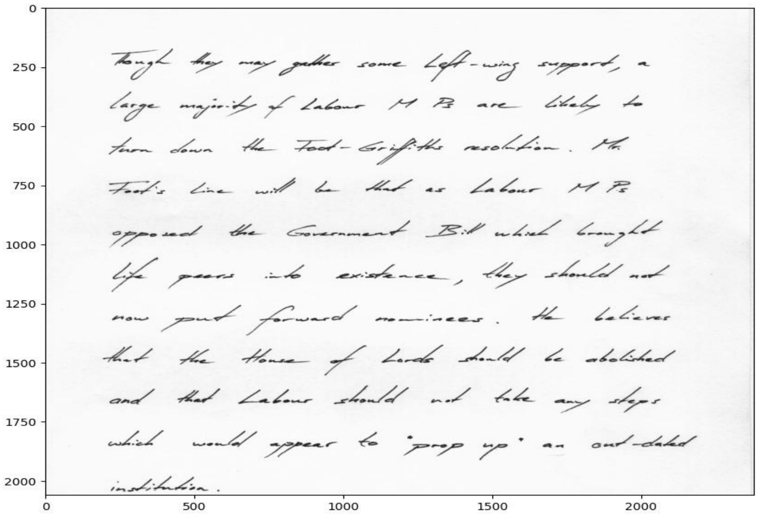
# PROPOSED METHODOLOGY

- Our project aims to develop an application utilizing deep learning concepts, particularly CNN, to achieve accurate handwriting recognition, addressing the growing demand for efficient data management solutions.
- Utilize data augmentation and domain adaptation to improve model performance on diverse handwriting styles and limited datasets.
- Design lightweight models and explore hardware acceleration for faster processing on mobile devices or resource-constrained environments

# ARCHITECTURE

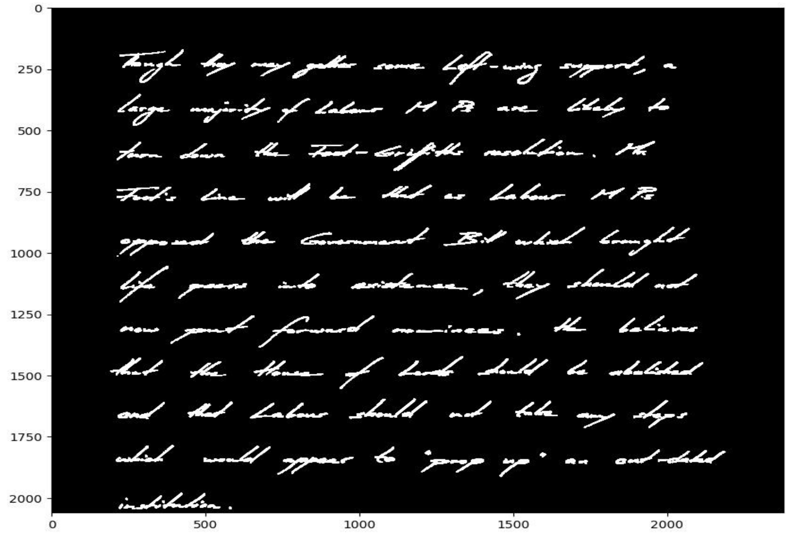


# IMPLIMENTATION

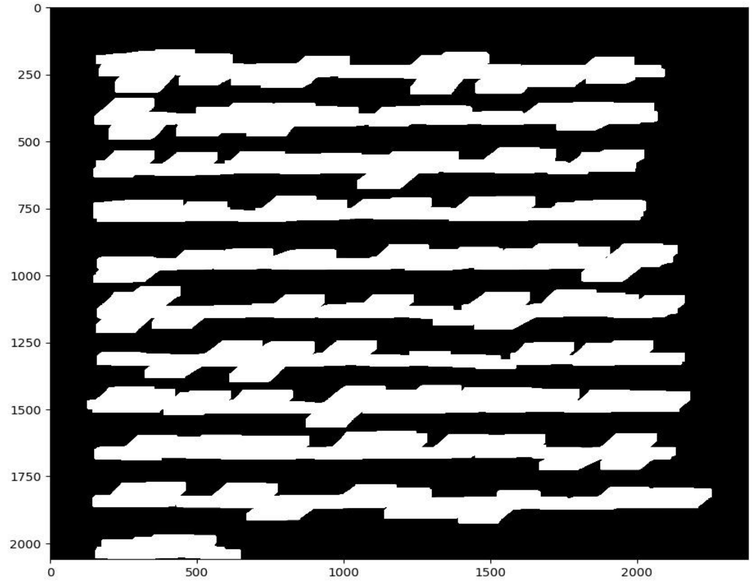

0
250
500
750
1000
1250
1500
1750
2000
0
500
1000
1500
2000

Though they may gather some Left-wing support, a
large majority of Labour M Ps are likely to
turn down the Foot-Griffiths resolution. Mr.
Foot's line will be that as Labour M Ps
opposed the Government Bill which brought
life peers into existence, they should not
now put forward nominees. He believes
that the House of Lords should be abolished
and that Labour should not take any steps
which would appear to "prop up" an out-dated
institution.

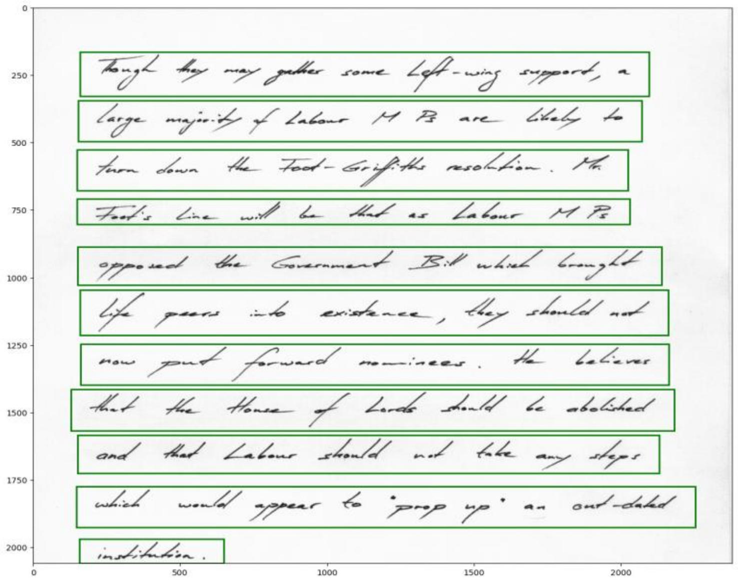
1.Sample data



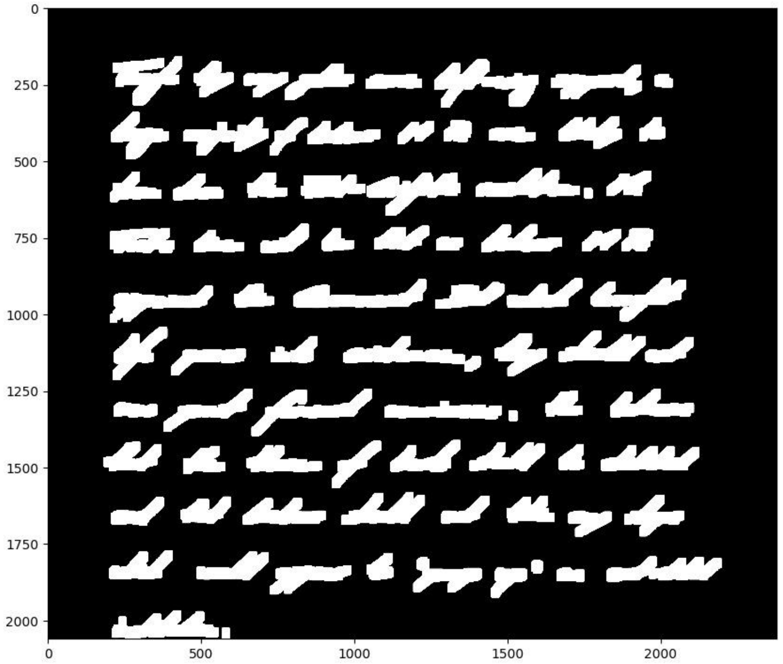
2.Grey Scale



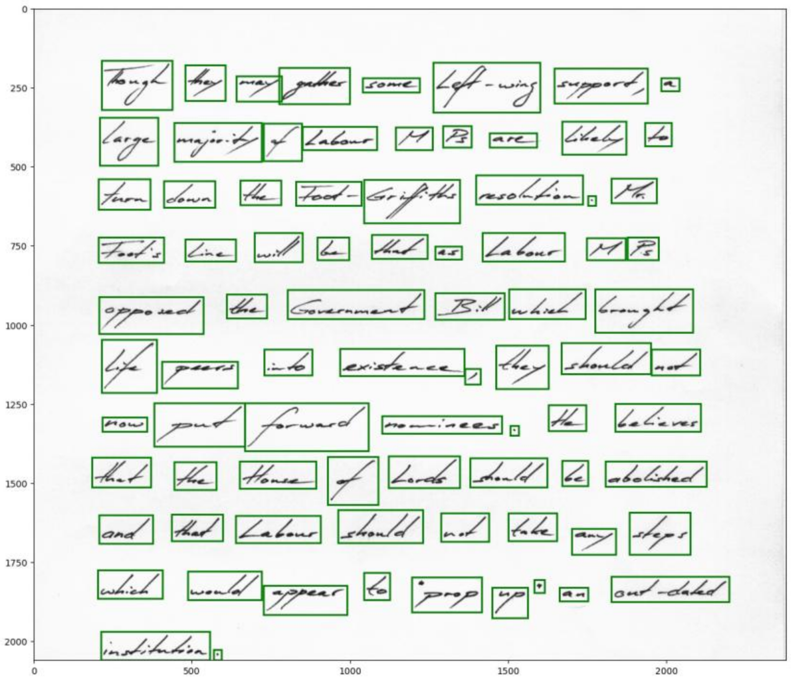
3.Dilation



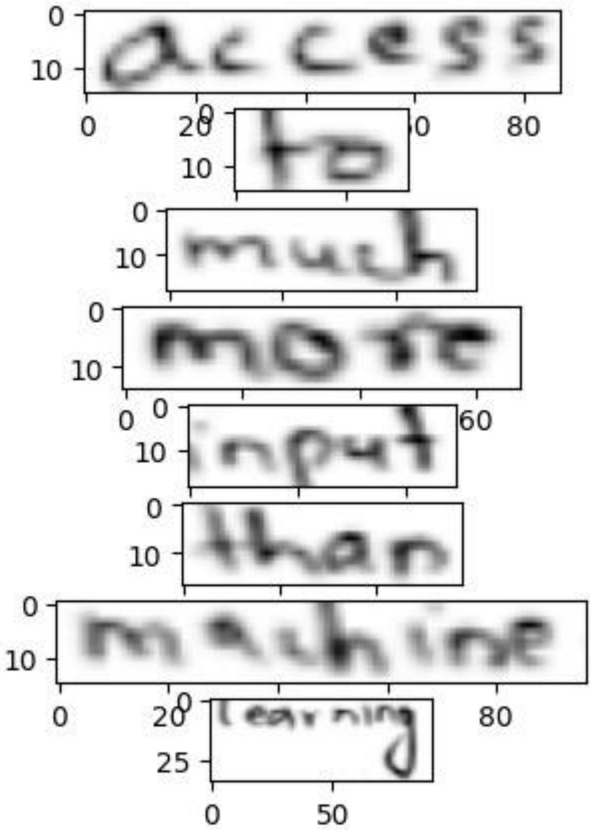
4.Find contours



5.Dilation

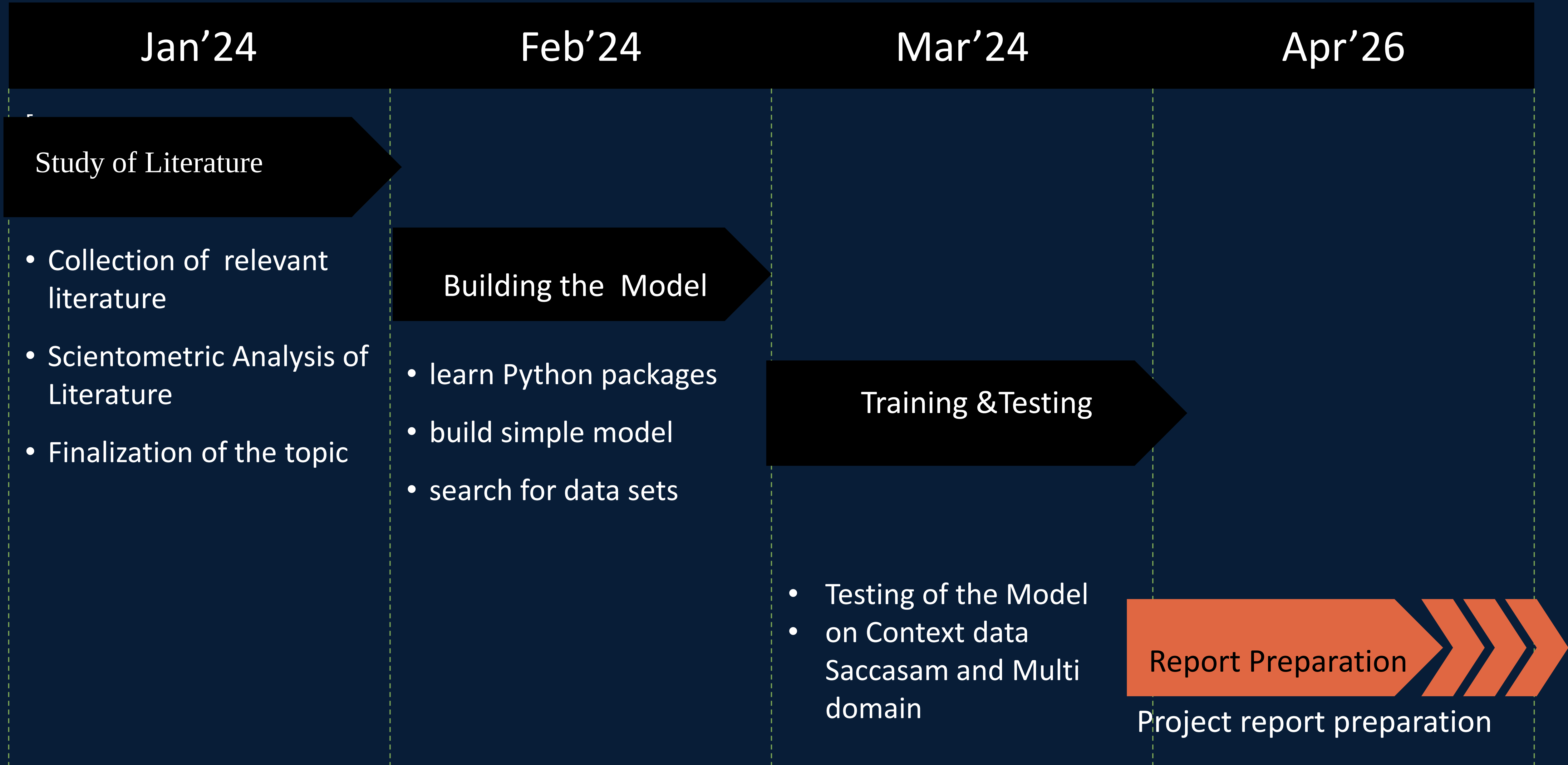


6. sort contours from left
to right and
top to bottom



7. scale to specified
height

# Timeline & Work Plan/ Progress



Jan’24

Feb’24

Mar’24

Apr’26

Study of Literature

- Collection of relevant literature
- Scientometric Analysis of Literature
- Finalization of the topic

Building the Model

- learn Python packages
- build simple model
- search for data sets

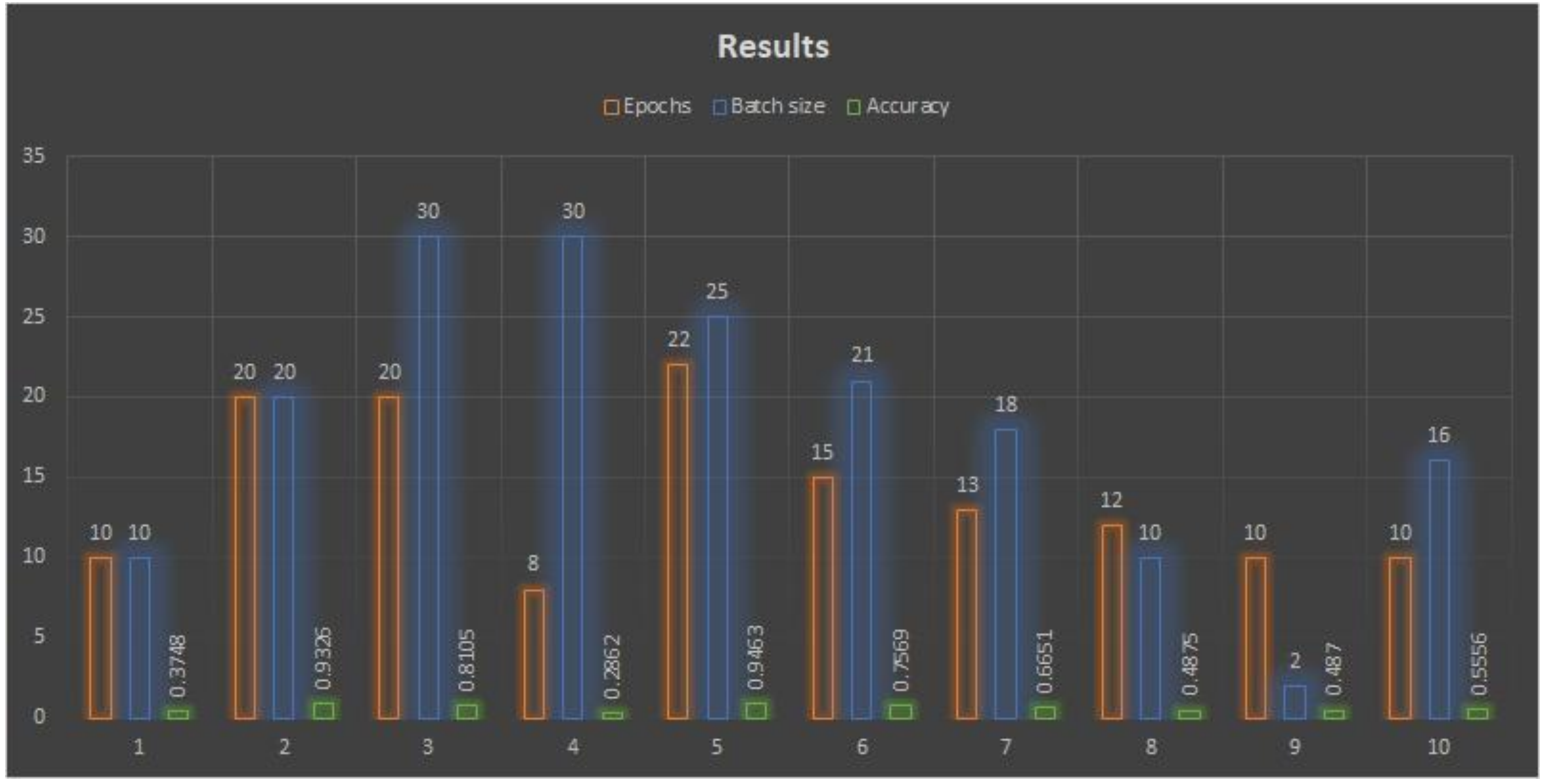
Training &Testing

- Testing of the Model
- on Context data Saccasam and Multi domain

Report Preparation

Project report preparation

# Results


Results
□ Epochs □ Batch size □ Accuracy
35
30
25
20
15
10
5
0
10 10 0.3748
20 20 0.9326
20 30 0.8105
8 30 0.2862
22 25 0.9463
15 21 0.7569
13 18 0.6651
12 10 0.4875
10 2 0.487
10 16 0.5556
1
2
3
4
5
6
7
8
9
10

# ...continued (results)

- **Impact of Epochs:** In general, validation accuracy tends to improve with an increased number of epochs, suggesting the model benefits from more training iterations. However, exceptions exist (e.g., 30 epochs vs. 20 epochs). This might indicate reaching a point of diminishing returns or potential overfitting with a very high number of epochs.
- **Impact of Batch Size:** There isn't a clear trend regarding the impact of batch size on validation accuracy. This suggests that within the tested range (2 to 30), batch size might not be a critical factor for your model's performance.
- **Best Configuration:** The best validation accuracy of 0.9463 was achieved with a batch size of 25 and 22 epochs.

# CONCLUSION AND FUTURE SCOPES

- **Exploration of Advanced Architectures:** Investigating more advanced architectures like attention mechanisms or deeper LSTM layers could potentially improve performance.
- **Adaptation to Different Datasets:** Adapting the model to recognize different writing styles or languages can broaden its applicability.
- **Real-World Application Integration:** Exploring the integration of this model into real-world applications, such as document analysis systems or handwriting recognition tools, can demonstrate its practical value.

SR UNIVERSITY

# Thank you

SR
UNIVERSITY