Subjective Answer Evaluator using CRNN and LSTM
An application of OCR and NLP

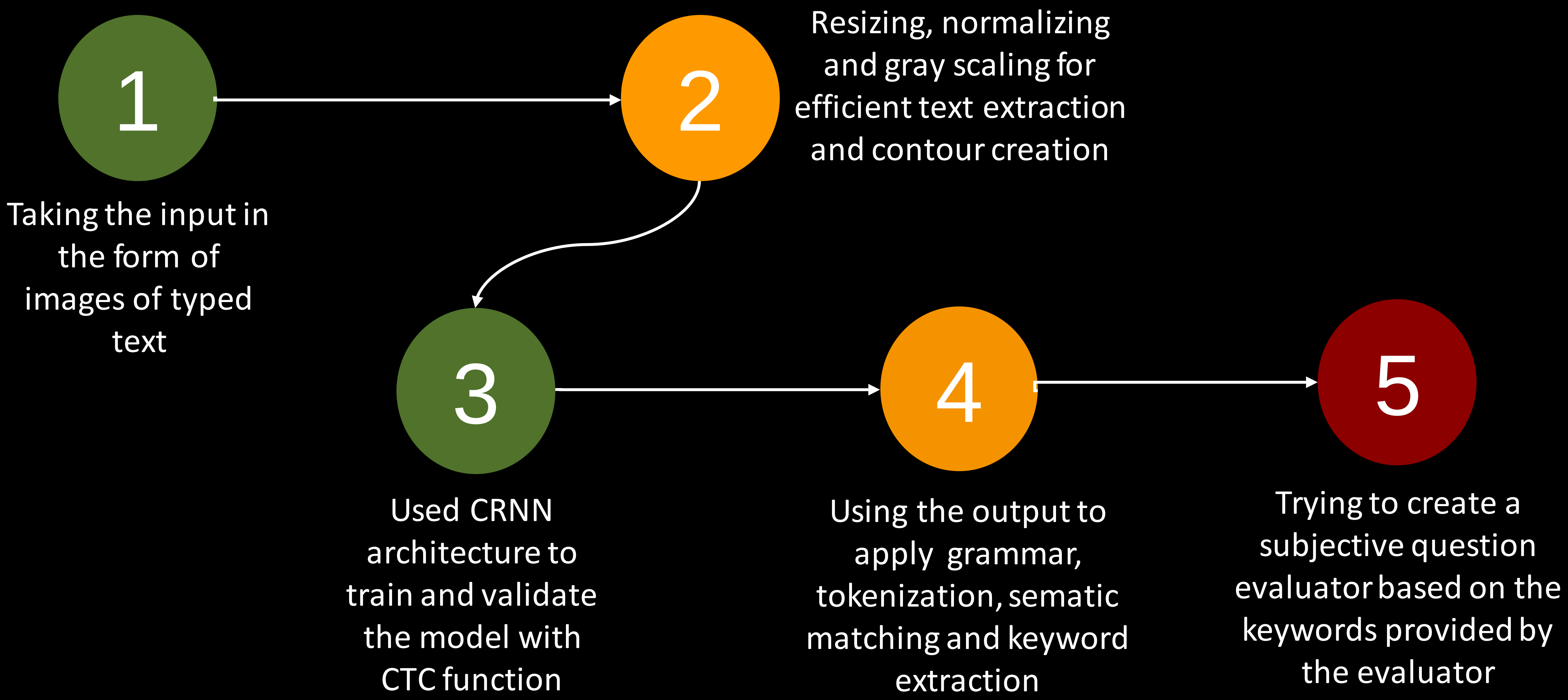
Literature Review

- **Text-recognition Models:**
  - Usage of Edge Detection techniques (contour formation) to retrieve data efficiently.
  - Using CRNN method for text-recognition with an accuracy of 70.4% instead of Hough Transform or pytesseract or CRAFT –text pretrained model.
  - Evaluated the factors that affect the choice of dataset and dividing the data for training, validation and testing efficiently such as :
    - What is the type of the dataset ?
    - If we use more images for training while creating a text-recognition model will the accuracy of the model changes?
    - Size of the dataset and its feasibility.
  - Using the CTC loss function as an efficient decoder using just the ground truth text and the output from the Neural Network Model.

Purpose of our Project

- To increase the accuracy of the OCR by applying neural network methodologies to the existing methodologies.
- By combining two processes of extracting the features from the text dataset and provide a clear pattern for the evaluator.
- To expand the scope of the model to other types of question evaluations and use the same model for feature/text extraction from financial and bank transcripts, healthcare transcripts etc.
- Provide a comparative analysis of our project with the existing pre-trained models.

Phases of our Project



1

Taking the input in the form of images of typed text

2

Resizing, normalizing and gray scaling for efficient text extraction and contour creation

3

Used CRNN architecture to train and validate the model with CTC function

4

Using the output to apply grammar, tokenization, sematic matching and keyword extraction

5

Trying to create a subjective question evaluator based on the keywords provided by the evaluator

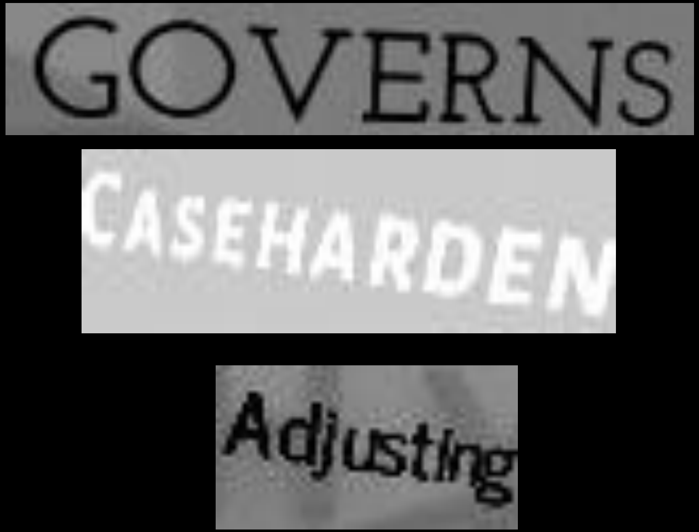
Dataset

Dataset used for the typed text:

- Dataset : https://www.robots.ox.ac.uk/~vgg/data/text/#sec-synth
- Author : Max Jaderberg, Karen Simonyan , Andrea Vedaldi and Andrew Zisserman
- Type : Synthetic Data And Artificial Neural Network for Natural Scene Text Recognition
- Organization : Visual Geometry Group, Department of Engineering Science, University of Oxford
- Dataset size : 9 million images that helps train over 90K words
- No. of images used for training and validating the model : 52K images
- No. of images used to test on the epochs : 10

Images from Dataset

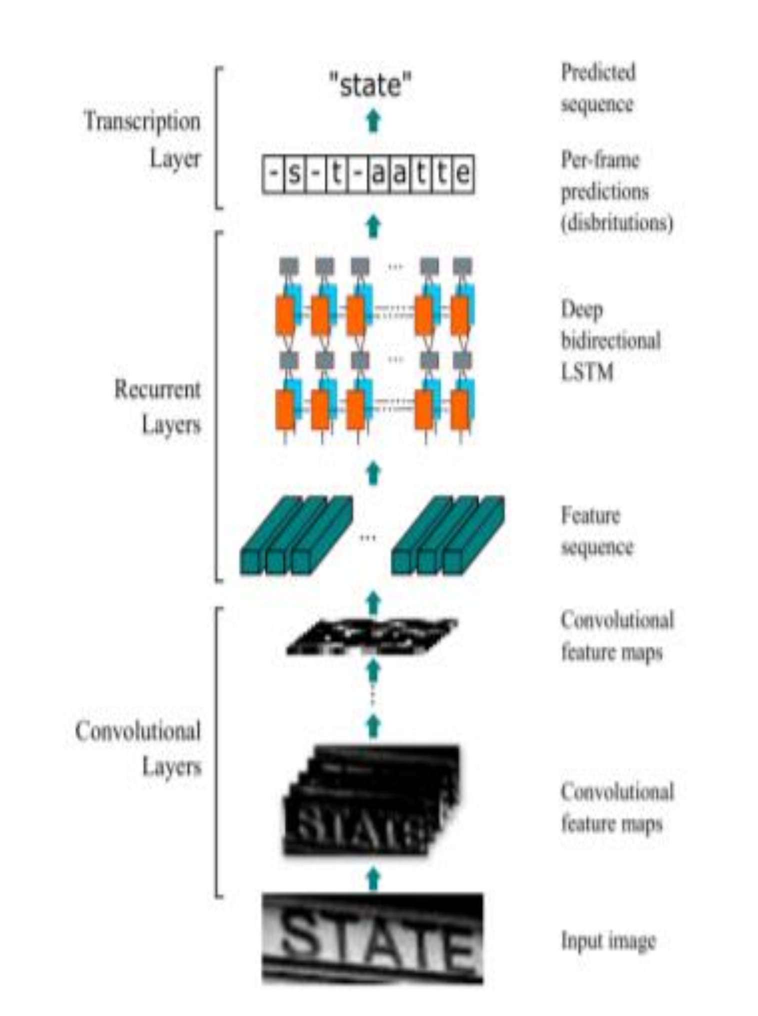
Sample Images from typed text database in word format



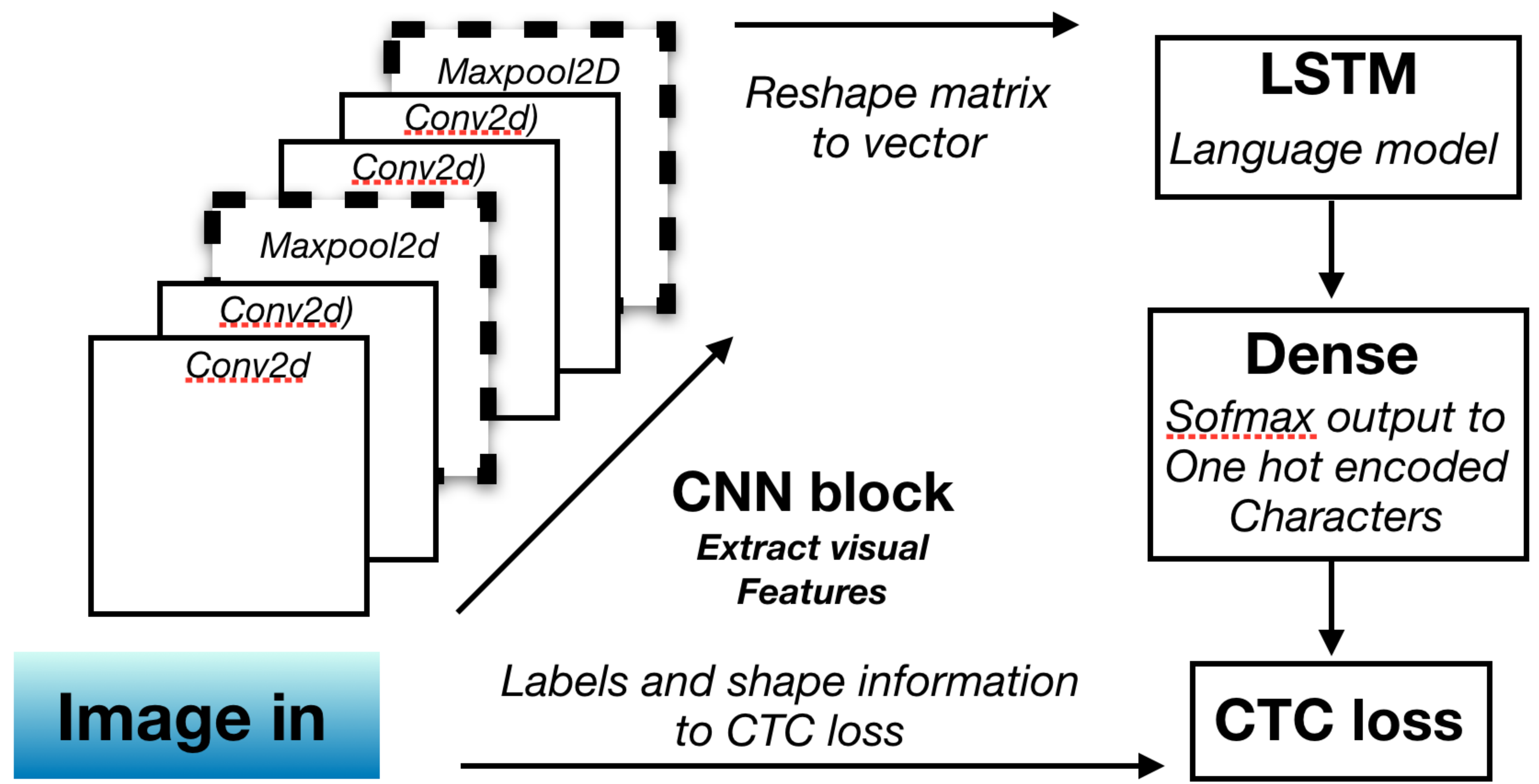
Phase 1 Image -Text Recognition

Steps to extract the text from the images :

- Creating a contour around words in the labeled sentences.
- Creating a 7-layer CNN Model with kernel size of 3x3 and neuron numbers ranging from 64 to 512.
- The Activation function used is ***ReLu*** to devoid the neural network of any exponential growth in computation.
- Apply batch normalization to 256 and feed the trained CNN to 2 layers of Bidirectional LSTM model.
- LSTM model provides **softmax output** with convolutional feature maps.
- These features maps are fed to CTC loss function to evaluate the loss of the data extracted.



Phase 1 Process Flow Diagram

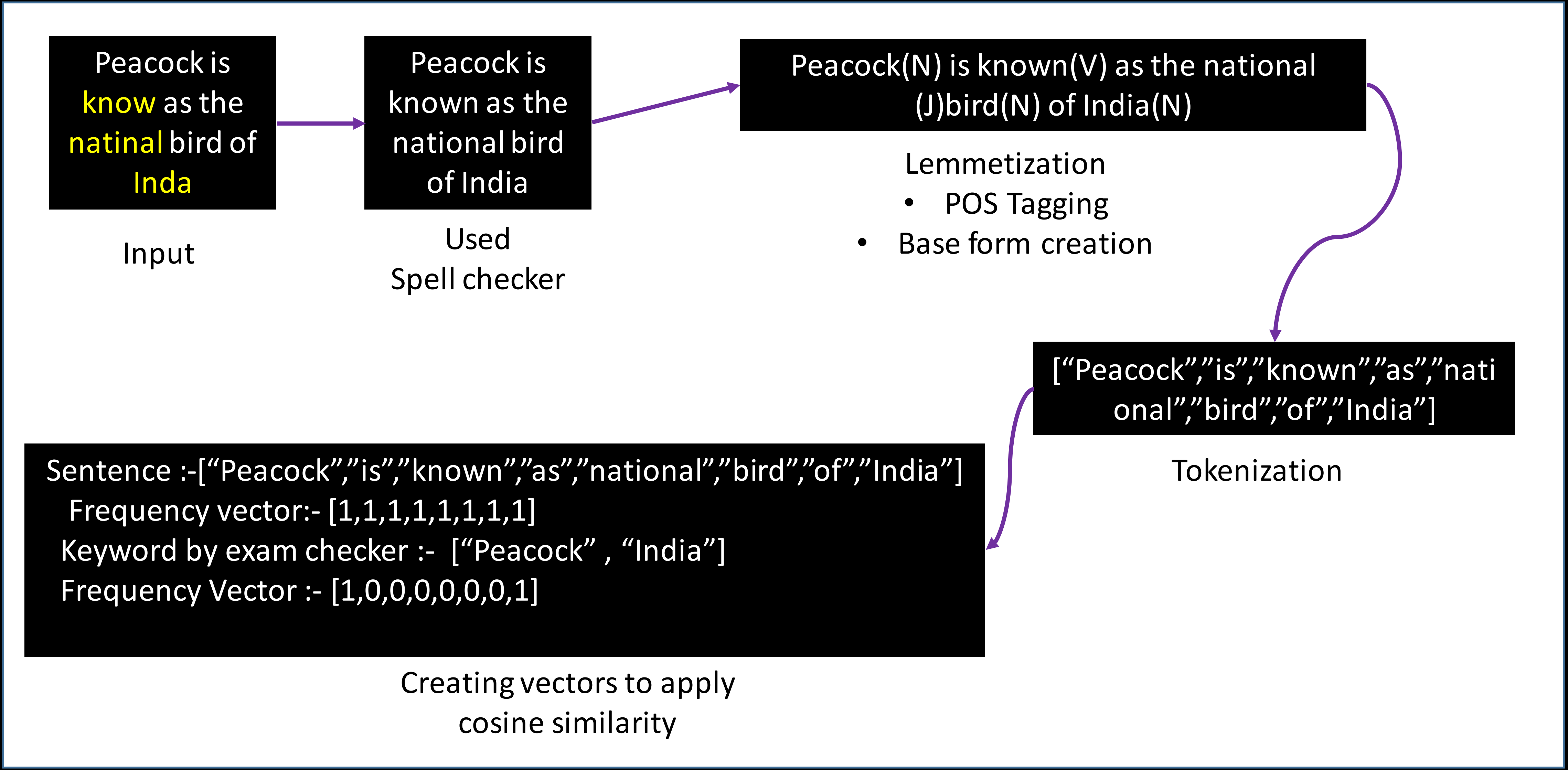


Phase 2 Answer Evaluation

Steps to create an evaluator the text from the image:

- Since this is the application of NLP, we have used **NTLK** Library for various purposes mentioned below.
- The output sentence that is obtained from the OCR will be fed into our NLP model.
- Step 1 : Spellchecker library performs spellchecking on the sentence received as output.
- Step 2: Text cleaning function is used to remove extra spaces, hyperlinks, regular expression etc.
- Step 3: After spellchecking and cleaning we implement the concept of lemmatization i.e. **POS Tagging and converting the words in the sentence to its base form.**
- Step 4: Now we apply the method of tokenization to store the sentence into words to perform the similarity function in the further stages.
- Step 5: To remove the frequent word such as **‘is,the,a,an etc.’** we have used **Stopword** function of the NLTK library for only English dialect.
- Step 6 : We have created vector of the output sentence and the sentence or key word entered by the exam checker
- Step 7: Applied the concept of **Cosine Simmilarity** to find the match in the answer.

Phase 2 : Process Flow



Peacock is know as the natinal bird of Inda

Input

Peacock is known as the national bird of India

Used
Spell checker

Peacock(N) is known(V) as the national (J)bird(N) of India(N)

Lemmetization
- POS Tagging
- Base form creation

[“Peacock”,”is”,”known”,”as”,”national”,”bird”,”of”,”India”]

Tokenization

Sentence :-[“Peacock”,”is”,”known”,”as”,”national”,”bird”,”of”,”India”]
Frequency vector:- [1,1,1,1,1,1,1,1]
Keyword by exam checker :- [“Peacock” , “India”]
Frequency Vector :- [1,0,0,0,0,0,0,1]

Creating vectors to apply
cosine similarity

Libraries Used

- **os module** : to remove, add and fetch directories from local device. Used to save the modal and create an OS independent API so that we don't have to train the model everytime we load the code
- **Fnmatch :** to keep a check , so that there are no duplicate entries of the directories
- **matplotlib :** to plot the CTC loss value and accuracy obtained.
- **keras :** for pre-processing of images and importing CNN and LTSM models
- **tensorflow :** to check the availability of GPU and calculate its run time and handle GPU allocation.
- **OpenCV – PyImageSearch :** for contouring the words from the sentences and applying padding properties to the images extracted.
- **NLTK :** for spellchecking, cleaning the text, tokenization.
- **Wordnetlemmetizer:** Used for POS Tagging and base form conversion

References and Citations

Papers:

1. https://www.acadpubl.eu/hub/2018-118-24/3/577.pdf
2. https://cdn.iiit.ac.in/cdn/cvit.iiit.ac.in/images/ConferencePapers/2019/PID6008523.pdf
3. http://www.ijsrp.org/research-paper-0315/ijsrp-p3919.pdf
4. https://www.ijcaonline.org/archives/volume179/number31/garg-2018-ijca-916390.pdf

Other documentation:

1. https://nanonets.com/blog/ocr-with-tesseract/#trainingtesseractoncustomdata
2. https://cdn.iiit.ac.in/cdn/cvit.iiit.ac.in/images/ConferencePapers/2019/PID6008523.pdf
3. https://www.ijcaonline.org/archives/volume179/number31/garg-2018-ijca-916390.pdf
4. https://wandb.ai/authors/text-recognition-crnn-ctc/reports/Text-Recognition-With-CRNN-CTC-Network--VmlldzoxNTI5NDI
5. https://medium.com/analytics-vidhya/image-text-recognition-738a368368f5
6. https://www.ijert.org/using-crnn-to-perform-ocr-over-forms
7. https://repositum.tuwien.at/retrieve/10807