### it llogo2.JPGREPORT GENERATION ON CHEST X-RAYS USING DEEP LEARNING

*PROJECT REPORT submitted in partial fulfillment of the requirements For the award of the degree*

**BACHELOR OF TECHNOLOGY**

**IN**

**INFORMATION TECHNOLOGY**

*Submitted by*

**M.J.N. VENKATA SAI 208W1A12A0 PANITINI MONICA 208W1A12A5 VEMULAPALLI SAIESH 208W1A12C7**

*Under the Guidance of*

**Dr. M. ASHOK KUMAR**

**Assistant Professor**



**DEPARTMENT OF INFORMATION TECHNOLOGY V R SIDDHARTHA ENGINEERING COLLEGE**

**(AUTONOMOUS - AFFILIATED TO JNTU-K, KAKINADA)**

#### Approved by AICTE &Accredited by NBA

**KANURU, VIJAYAWADA-520007 ACADEMIC YEAR**

**(2022-23)**

# V.R. SIDDHARTHA ENGINEERING COLLEGE

(Affiliated to JNTUK: Kakinada, Approved by AICTE, Autonomous) (An ISO certified and NBA accredited institution)

Kanuru, Vijayawada – 520007



## CERTIFICATE

This is to certify that this project report titled **“Report Generation on chest x-rays using deep learning”** is a bonafide record of work done by **M.J.N. Venkata Sai (208W1A12A0), Panitini Monica (208W1A12A5), Vemulapalli Saiesh (208W1A12C7),** under my guidance and supervision is submitted in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Information Technology, **V.R. Siddhartha Engineering College** (Autonomous under JNTUK) during the year **2022-2023**.

###### Dr. M. Ashok Kumar Dr. M. Suneetha

Assistant Professor Professor & Head

Dept. of Information Technology Dept. of Information Technology

**EXTERNAL EXAMINER SIGNATURE**

*Date of examination*

### TABLE OF CONTENTS

[ACKNOWLEDGEMENT iii](#_TOC_250015)

[LIST OF FIGURES iv](#_TOC_250014)

[LIST OF TABLES v](#_TOC_250013)

LIST OF SYMBOLS vi

[ABSTRACT vii](#_TOC_250012)

CHAPTER-1 Introduction 1

* 1. [Origin of the Problem 1](#_TOC_250011)
  2. [Basic definitions and Background 2](#_TOC_250010)
  3. [Problem Statement with Objectives and Outcomes 5](#_TOC_250009)
  4. Societal Applications of Proposed work 6

[CHAPTER-2 Review of Literature 7](#_TOC_250008)

* 1. [Description of Existing Systems 7](#_TOC_250007)
  2. [Summary of Literature Study 9](#_TOC_250006)
  3. Software Requirement Specification 9

CHAPTER-3 Proposed Method 10

* 1. [Design Methodology 10](#_TOC_250005)
  2. [System Architecture Diagram 11](#_TOC_250004)
  3. Description of Algorithms 11
  4. [Description of datasets, Requirements and Tools 12](#_TOC_250003)

CHAPTER-4 Results and Observations 14

* 1. [Stepwise description of Results 14](#_TOC_250002)
  2. Test case results and Result Analysis 15
  3. Observations from the work 16

[CHAPTER-5 Conclusion and Future work 19](#_TOC_250001)

* 1. Conclusion 19
  2. Future study 19

[References 20](#_TOC_250000)

### ACKNOWLEDGEMENT

First and foremost, I sincerely salute our esteemed institution **V.R SIDDHARTHA ENGINEERING COLLEGE** for giving me this opportunity for fulfilling my project. I am grateful to our principal **Dr. A.V.RATNA PRASAD**, for his encouragement and support all through the way of my project.

On the submission of this Project report, I would like to extend my honour to **Dr. M. Suneetha**, Head of the Department, IT for her constant motivation and support during the course of my work.

I feel glad to express my deep sense of gratefulness to my project guide **Dr. M. Ashok kumar, Assistant Professor** for his guidance and assistance in completing this project successfully.

I would also like to convey my sincere indebtedness to all faculty members, including supporting staff of the Department, friends and family members who bestowed their great effort and guidance at appropriate times without which it would have been very difficult on my part to finish the project work.

### LIST OF FIGURES

**Figure Page**

Figure 1.1: Image Captioning 1

Figure 1.2: XML report 2

Figure 1.3: Chexnet Model 5

Figure 3.1: Flow Diagram 10

Figure 3.2: System Architecture Diagram 11

Figure 3.3: LSTM 11

Figure 3.4: GRU 12

Figure 3.5: Dataset 12

Figure 4.1: CheXnet model 14

Figure 4.2: Training the model 14

Figure 4.3: Evaluation model 15

Figure 4.4:Testcase1 15

Figure 4.5:Testcase2 15

Figure 4.6:Testcase3 16

Figure 4.7:Images in a report 16

Figure 4.8:Word count distribution 17

Figure 4.9: Word cloud 17

Figure 4.10:Train vs validation loss graph 18

### LIST OF TABLES

#### Tables Page

Table 2.1 Literature survey 7

### LIST OF FORMULAS

#### Formulas Page

Formula 1.1 BLEU formula 3

Formula 1.2 BP formula 4

### ABSTRACT

The chest x-ray is the most commonly performed diagnostic x-ray examination. A chest x- ray produces images of the heart, lungs, airways, blood vessels, and the bones of the spine and chest. Often time, it is the duty of a radiologist to conclude these x-rays so that to give appropriate treatment to the patients. It is often time-consuming and tedious to get detailed medical reports from these x-rays. In high population countries, a radiologist may come across 100s of x-ray images. This project aims to present a compilation of the most outstanding deep learning strategies focused on the automatic generation of medical reports from X-Ray images. In order to handle this challenging problem, deep learning algorithms have been include with models, to get promising results. So if a properly learned deep learning model can automatically generate these medical reports, considerable work and time can be saved. In this project an Encoder and decoder with attention model is used to generate the text report and pretrained CheXnet model is used to obtain the image features. The generated text report is evaluated by BLEU score.

**Keywords**: Chest x-rays, Radiologist, Encoder, Decoder, Attention, BLEU, Chexnet

### CHAPTER-1

#### Introduction

This chapter summarizes the project's goal, origins, and applications. It also explains the project's need and scope, as well as the project's roadblocks, which provides useful information. The specifics of how and where the project can be used

#### Origin of the problem:

Image captioning is one of the most important and challenging tasks in deep learning. It is the process of generation of a textual description for an image. From the images, we can come up with a caption or description. This is the main idea behind the project And so far many deep learning models are achieved to perform this task. The model understands the contents of the image and generates the corresponding textual description. For example



Fig 1.1 Image Captioning

The applications of image captioning include Google Image search, **medical report generation**, etc.

It is very helpful in medical field where Radiologists need to describe medical X-ray images. Summarizing the X-ray in a form of radiology report is a complex task and more care should be taken in generating reports. The radiology report is a complete study of X-ray images describing normal and abnormal conditions which makes the right decision to take proper medication. Hence, the Radiologist is expected to diligently summarize a report and in the process of writing medical reports usually takes around 5–10 minutes per report but the problem here is in a day the doctors have to write medical reports that number in 100s which can take a lot of their time and For less-experienced

radiologists and pathologists, especially those working in the rural areas where the quality of healthcare is relatively low, writing medical-imaging reports is demanding or on the other hand for experienced radiologists and pathologists, writing imaging reports can be tedious and time consuming so the objective of this project is to build a deep learning model that automatically generates the text description chest X-rays so that it can reduce the doctors’ time and reduces the some of the burden of the medical professional and here we taking a publicly available dataset from Indiana University which consists of chest X-ray images and reports (in XML format) which contain information regarding the findings and impression of the X-ray. The goal is to predict the impressions of the medical report attached to the images.

#### Basic definitions and Background:

* + 1. **Parsing XML file:**

The data consists of a set of x-ray images and XML files containing the medical report. As shown in given figure, this XML has a lot of information like the image id of the x- ray, indication, findings, impression, etc. We will extract the findings and impressions from these files and consider them as reports because they are more useful for the medical report. We also need to extract the image id from these files to get the x-rays corresponding to each report.

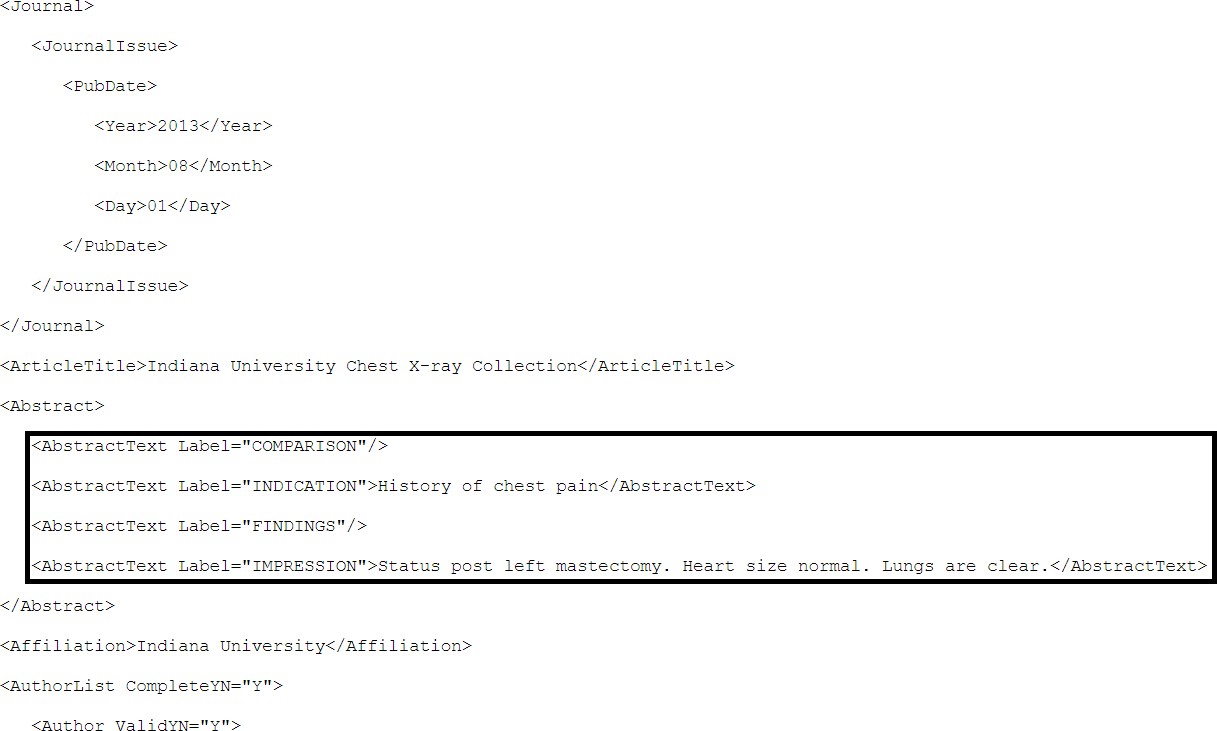


Fig :1.2. XML report

#### Pre-processing :

In this phase the text data are preprocessed to remove unwanted tags, texts,

punctuation and numbers. Perform basic decontractions i.e words like won’t,

can’t and so on will be converted to will not, can not and so on respectively. We will also check for the empty cell or NaN values and If there are any empty cells in the image name column we will drop those cells.Each text column word counts are calculated and added to the data frame column.If there any empty

or NaN value in text data we will replace it with “No <Column Name>” (ex: No Impression) After the data preprocessing step, we have a total of 3851 rows present in the final data points.

* + 1. **Structured Data:**

There are only two image types — Front and Lateral, but each patient has multiple x-rays associated with them. The maximum number of images associated with a report can be 5 while the minimum is 0. The highest frequency of being associated with a report are 2 images. We have more than 2 images and in some cases less than 2 images are associated with each data point. we need to come up with an idea to structure data point that can limit the data point to 2 images per data point For example

if I have 5 images then total 4 data points created

* + - * 1st image + 5th image
      * 2nd image + 5th image
      * 3rd image + 5th image
      * 4th image + 5th image

1. If I have 4 images then total 3 data points created

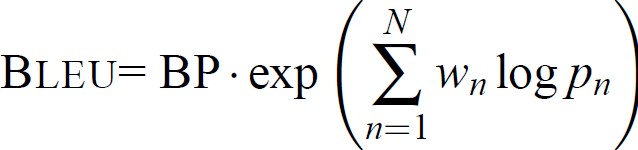
* 1st image + 4th image
* 2nd image + 4th image
* 3rd image + 4th image

1. If I have 3 images then total 2 data points created

* 1st image + 3rd image
* 2nd image + 3rd image

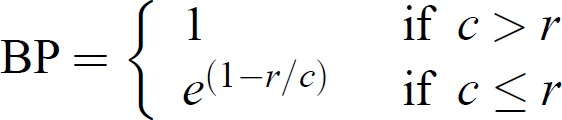
1. If I have 2 images then it is according to our requirement
2. At last, If i have 1 image, just replicate it and make it 2.
   * 1. **Evaluation metric(BLEU score):**

BLEU score stands for Bilingual Evaluation Understudy. Here we will be using BLEU score as the metric. BLEU score compares each word in the predicted sentence and compare it to the reference sentence (It is also done in n-grams) and returns score based on how many words were predicted that were in the original sentence. It returns a value between 0 and 1. The metric close to 1 means that the two are very similar.



Formula : 1.1 BLEU formula

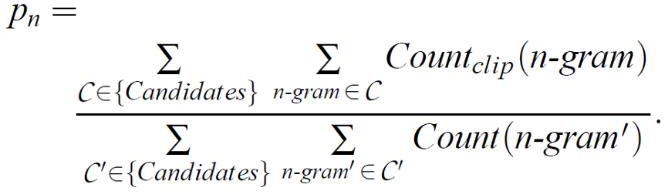
BP-brevity penalty



Formula : 1.2 BP formula r-count of words in reference translation

c- count of words in a candidate translation

N: No. of n-grams, we usually use uni-gram, bi-gram, 3-gram, 4-gram Wn: Weight for each modified precision, by default N is 4, Wn is 1/4=0.25 Pn: Modified precision



Formula 1.3 modified precision

#### Tokenization

When working with text, the first thing we must do come up with a strategy to convert strings to numbers (or to “vectorize” the text) before feeding it to the model, We will convert text data into numerical data using Tokenizer. The tensorflow deep learning library provides tools to perform this operation.

* + 1. **Transfer learning:**

Images along with partial reports are the inputs to our model. We need to convert every image into a fixed sized vector which can then be fed as input to the model. We will use **transfer learning** for this purpose.

CheXNet, is a 121-layer convolutional neural network trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset, containing over 100,000 frontal-view X-ray images with 14 diseases. However, our purpose here is not to classify the images but just to get the bottleneck features for each image. Therefore the last classification layer of this network is not needed.

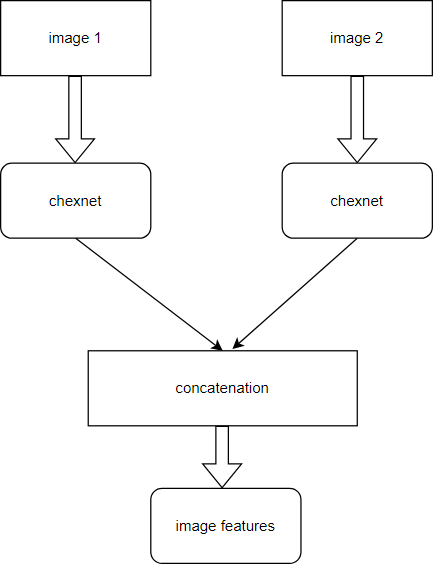


Figure 1.3 ChXnet Model

#### Encoder-Decoder architecture :

A sequence-to-sequence model is a deep learning model that takes a sequence of items (in our case, features of an image) and outputs another sequence of items (reports).The encoder processes each item in the input sequence, it compiles the information it captures into a vector called the context. After processing the entire input sequence, the encoder sends the context over to the decoder, The decoder calls the one-step attention layer for each of the decoder time-steps and computes the scores and attention-weights. All the outputs of each time-steps are stored in the ‘all-outputs’ variable. The outputs from each decoder step are the next word in the sequence. ‘all-outputs’ will be our final output.

#### Problem statement with objectives and outcomes:

The chest x-ray is the most commonly performed diagnostic x-ray examination. A chest x-ray produces images of the heart, lungs, airways, blood vessels, and the bones of the spine and chest. Often time, it is the duty of a radiologist to conclude these x-rays so that to give appropriate treatment to the patients. It is often time-consuming and tedious to get detailed medical reports from these x-rays. So if a properly learned machine learning model can automatically generate these medical reports, considerable work and time can be saved.

This project aims to present a compilation of the most outstanding deep learning strategies focused on the automatic generation of medical reports from X-Ray images. In order to handle this challenging problem, deep learning algorithms have been include with models, to get promising results.The medical report generated from the model in the final stage.

#### Realtime applications of proposed work:

Medical report generation task, which targets to produce long and coherent descriptions of medical images and as medical report generation is a challenging task since it is time- consuming and requires expertise from experienced radiologists. The goal of medical report generation is to accurately capture and describe the image findings. As the proposed work automatically generates the text report , it reduces the time for radiologists.

### CHAPTER- 2 REVIEW OF LITERATURE

This chapter mainly focuses on the resources that helped us to grasp an ideology of generating medical report on chest x-rays using deep learning. The research papers gave us the detail of how to develop a model that predicts medical report and how to apply analyzing different algorithms on the data.

#### Description of existing systems

This section mainly focuses on the research papers including the main details ofthe paper. It includes the title, authors and descriptions drawn from each paper.

Table [2.1]: Literature Survey of report generation on chest xrays using deep learning

|  |  |  |
| --- | --- | --- |
| **S.no** | **Title** | **Authors** |
| 1 | Attention based automated radiology report generation using CNN and LSTM | Mehreen SirsharID1 , Muhammad Faheem Khalil ParachaID1 \*, Muhammad Usman Akram1 , Norah Saleh Alghamdi2 , Syeda Zainab  Yousuf Zaidi3 , Tatheer Fatima |
| **Description:** In this paper, to address this issue an approach is proposed,based on the continuous integration of convolutional neural networks and long short-term memory for detecting diseases, followed by the attention mechanism for sequence generation  based on these diseases | | |
| 2 | Report Generation of Lungs Diseases from Chest X-ray using NLP | Iqra Naz1 , Shagufta Iftikhar1  , Anmol Zahra1 , Syeda Zainab  Yousuf Zaidi1 |
| **Description:** In their experiments, they described a strategy on the basis of CNN-  RNN architecture with attention mechanism. Predicted the results of their model trained on the dataset | | |

|  |  |  |
| --- | --- | --- |
| 3 | An Automated medical report generation | Rahul Sai,Sharmila Banu Kather,B.K.Tripathy |
| **Description**: In this paper, it mainly focuses on the models with a single LSTM decoder perform much worse than those with a hierarchical LSTM decoder | | |
| **S.no** | **Title** | **Authors** |
| 4 | Automated radiology report generation using conditioned transformers | [OmarAlfarghalyaRanaKhaledb](https://www.sciencedirect.com/science/article/pii/S2352914821000472#!) [AbeerElkoranyaMahaHelalbAl](https://www.sciencedirect.com/science/article/pii/S2352914821000472#!) [yFahmya](https://www.sciencedirect.com/science/article/pii/S2352914821000472#!) |
| **Description:** In this paper,they proposed a new deep learning that uses visual and semantic features to condition a pre-trained transformer and then add semantic similarity metrics besides word-overlap metrics for the quantitative analysis | | |
| 5 | Automatic Report Generation for Chest X- Ray Images via Adversarial Reinforcement Learning | [DaibingHou](https://ieeexplore.ieee.org/author/37088653605); [ZijianZhao](https://ieeexplore.ieee.org/author/37085550655); [YuyingLiu](https://ieeexplore.ieee.org/author/37088395711);  [Faliang Chang](https://ieeexplore.ieee.org/author/37530881600) |
| **Description:** In this paper, a novel medical report generation framework is proposed that considers both language fluency and diagnostic fluency. | | |
| 6 | Contrasive Attention for Automatic Chest X- ray Report Generation | [Xuewei Ma](https://arxiv.org/search/cs?searchtype=author&query=Ma%2C%2BX), [Fenglin](https://arxiv.org/search/cs?searchtype=author&query=Liu%2C%2BF)  [Liu,](https://arxiv.org/search/cs?searchtype=author&query=Liu%2C%2BF) [Changchang Yin](https://arxiv.org/search/cs?searchtype=author&query=Yin%2C%2BC), [Xian](https://arxiv.org/search/cs?searchtype=author&query=Wu%2C%2BX) [Wu,](https://arxiv.org/search/cs?searchtype=author&query=Wu%2C%2BX) [Shen Ge](https://arxiv.org/search/cs?searchtype=author&query=Ge%2C%2BS), [Yuexian](https://arxiv.org/search/cs?searchtype=author&query=Zou%2C%2BY) [Zou,](https://arxiv.org/search/cs?searchtype=author&query=Zou%2C%2BY) [Ping Zhang](https://arxiv.org/search/cs?searchtype=author&query=Zhang%2C%2BP), [Xu Sun](https://arxiv.org/search/cs?searchtype=author&query=Sun%2C%2BX) |
| **Description**: In this paper, to effectively capture and describe abnormal regions they proposed an contractive attention model | | |
|  | | |

#### Summary of Literature Study

In one of the above papers suggested a way of an application to create automated textual reports for CXR, with the aim of assisting medical professionals in creating reports more efficiently and effectively. It is based on a CNN feature extraction model that acts as an encoder that converts an image into a fixed act as an encoder that converts an image into a fixed-size vector representation, followed by an RNN decoder that generates corresponding sentences based on the learned image features.

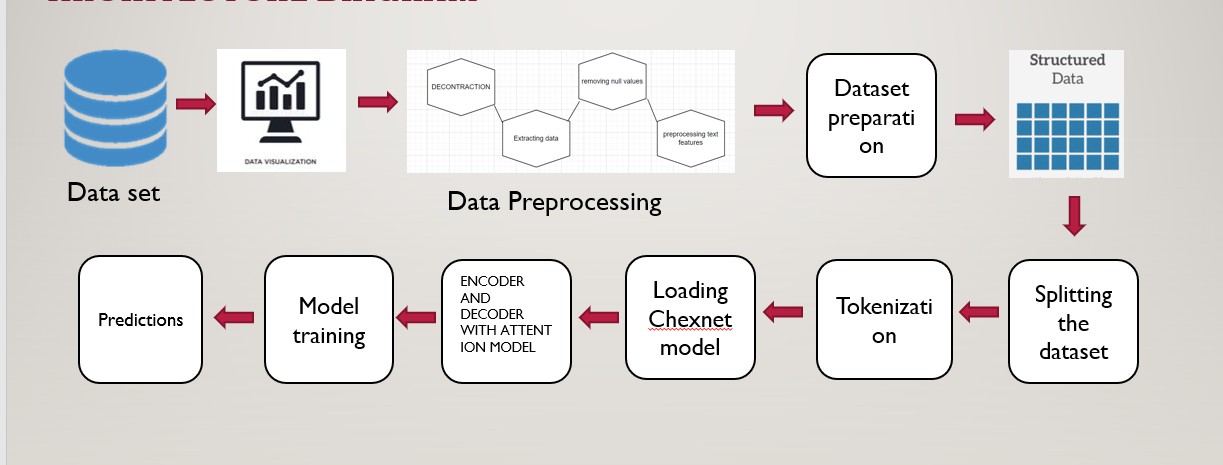
* 1. **Software Requirements Specifications:**
     + Python
     + Tensorflow
     + Keras
     + Python libraries

### CHAPTER-3 DESIGN METHODOLOGY

This chapter focuses on procedure for developing the deep learning model for generating the text report on chest x-rays.

#### Design Methodology:

In this project , first data set is collected from the Indiana university which contains the chest x-rays and xml reports there are around 7471 x-ray images and 3955 xml reports then the sample data are visualized. Next the data from the xml reports are extracted using python libraries and then it goes to pre-processing next creating a csv file which is our new dataset and need it need to be split into train ,validation and test data.Then we need a structured data for loading two images per patient and further train data is tokenized and here we are loading a chexnet model which is a pre trained model over 14000 images with accuracy 92.03% and the then writing encoder and decoder with attention model encode here is LSTM and decoder is GRU. And model is trained on validation dataset an custom loss function is written and finally made predictions on test data.



#### Implementation steps

Fig: 3.1 Flow diagram

* IMPORTING LIBRARIES AND DATASETS (BOTH X-RAYS AND XML REPORTS)
* VISUALIZING THE SAMPLE DATA
* STRUCTURED DATA
* DATA PREPROCESSING
  + TEXT DECONTRACTION
  + EXTRACTING THE DATA FROM XML FILES
  + REMOVING THE NULL VALUES
  + PREPROCESSING THE TEXT FEATURES
    - DISPLAYING SAMPLE IMAGES WITH TEXT FEATURES
    - DATASET PREPARATION
* SPLITTING THE DATASET
* TOKENIZATION
* LOAD CHXNET MODEL
* ENCODER AND DECODER WITH ATTENTION MODEL
* CUSTOM LOSS FUNCTION
* MODEL TRAINING
* PREDICTIONS ON SAMPLE TEST DATA

#### System Architecture Diagram:

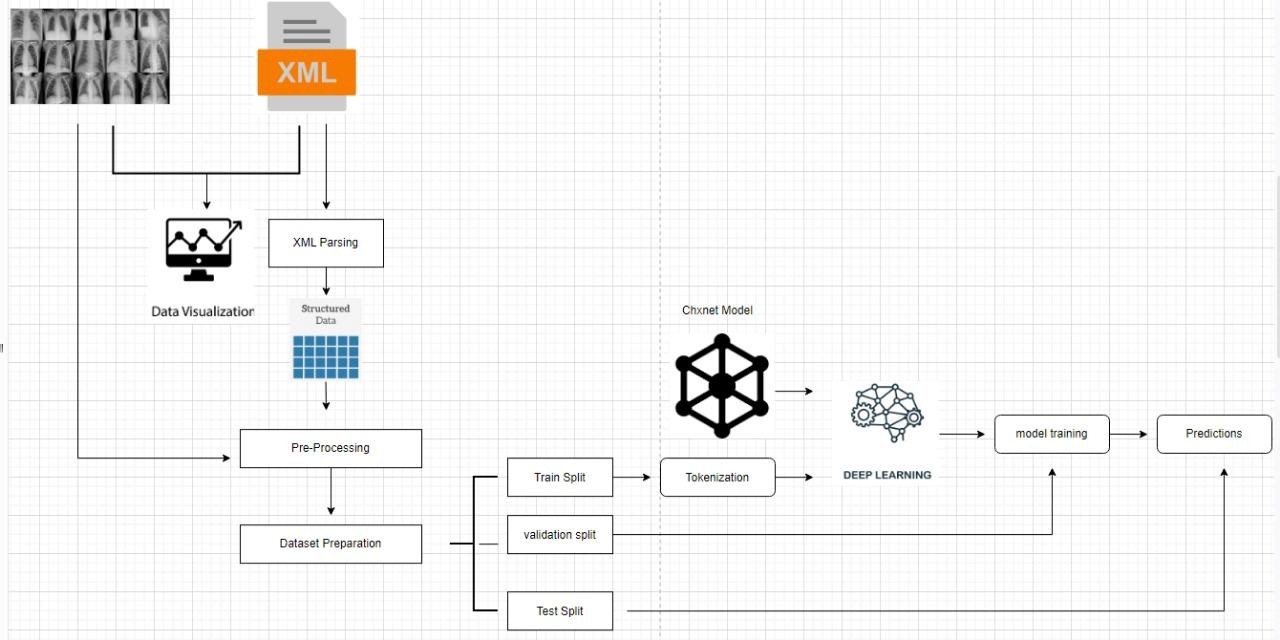


Fig: 3.2 System Architecture diagram

#### Description of the Algorithm:

**Algorithm-LSTM**

1. **Forget Gate(f) :**It determines to what extent to forget the previous data.
2. **Input Gate(I):**It determines the extent of information be written onto the Internal Cell State.
3. **Input Modulation gate(g):**It is considered as a sub-part of the input gate. It is used to modulate the information that the Input gate will write onto the Internal State Cell
4. **Output Gate(o):**It determines what output(next Hidden State) to generate from the current Internal Cell State.

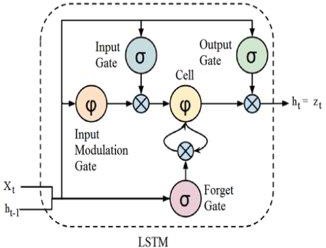


Fig : 3.3 LSTM

#### Algorithm – GRU

###### Update Gate(z) :

It determines how much of the past knowledge needs to be passed along into the future. It is analogous to the Output Gate in an LSTM recurrent unit.

###### Reset Gate(r):

It determines how much of the past knowledge to forget.

###### Current memory Gate(h(t)):

it is a sub-part of the Reset gate and to reduce the effect that previous information has on the current information that is being passed into the future.

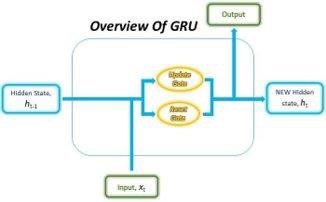


Fig: 3.4 GRU

#### Description of datasets, Requirements and Tools:

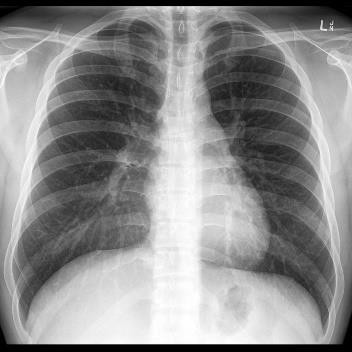
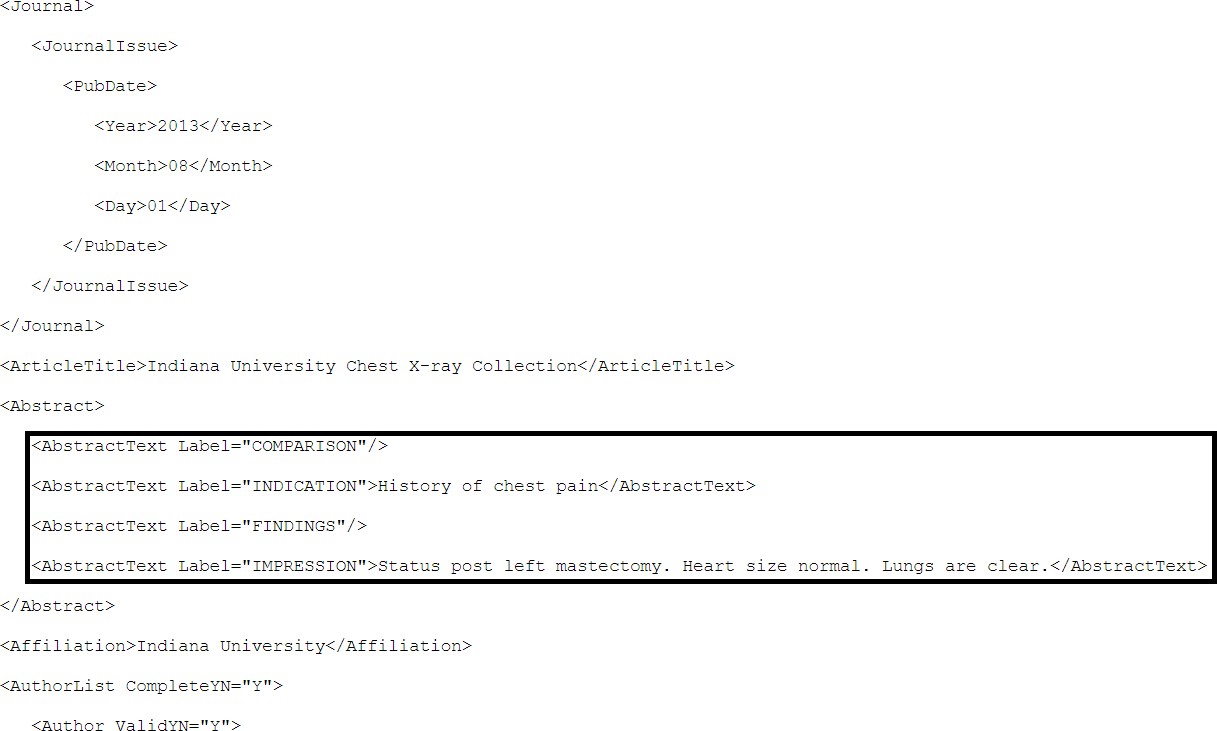
* + 1. **Data set**

Fig: 3. Dataset

The data for this problem is provided by the Indiana University hospital network. The data contains two parts. The X-rays contain a set of chest x-ray images . The reports contain the medical report of an x-ray. Total X-ray images are 7471 and xml reports are 3955.

#### User Interface:

This system's user interface is the Google Colab, which is a user-friendly Python Graphical User Interface.

###### Hardware Interfaces:

Python capabilities are used to allow the user to interact with the console.

###### Software Interfaces:

Required modules (ChexNet model) have been imported into the Python environment.

###### Hardware Requirements:

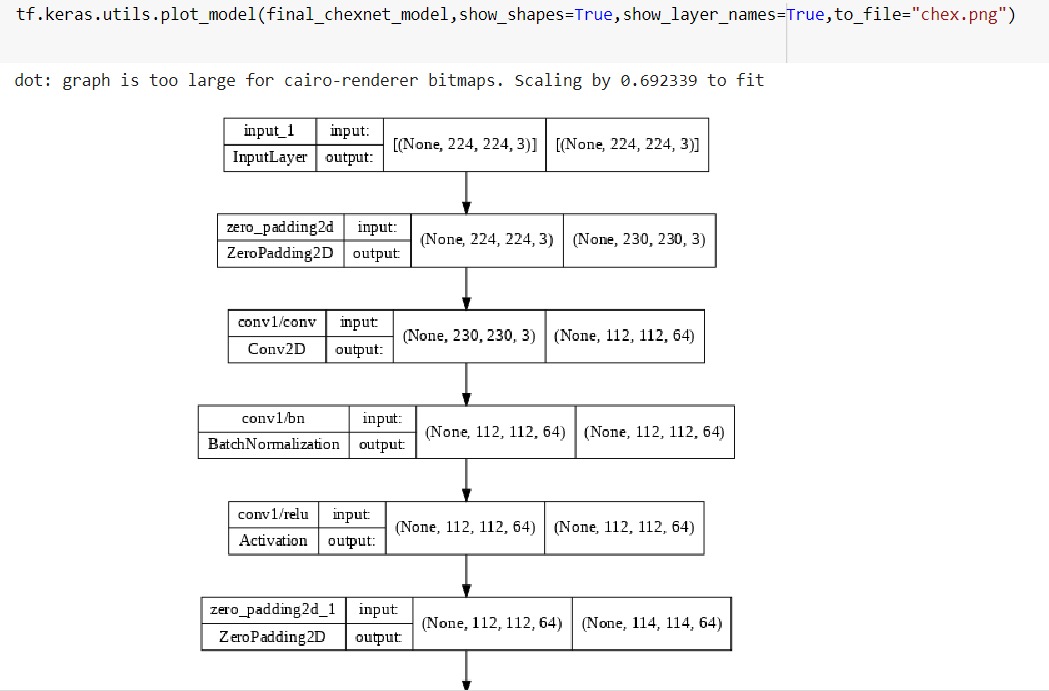
* + - 1. Processor – Pentium-IV
      2. RAM – 4GB (Minimum)
      3. HDD/SSD – 256GB (Minimum)

### CHAPTER-4

**Results & Observations**

#### Stepwise description of Results:

**A Loading the pre trained cheXnet model**



##### Figure 4.1 Chexnet Model

The figure 4.1.1 depicts the pretrained chexnet model, CheXNet, is a 121- layer convolutional neural network trained on ChestX-ray14, currently the largest publicly available chest X-ray dataset, containing over 100,000 frontal-view X-ray images with 14 diseases.

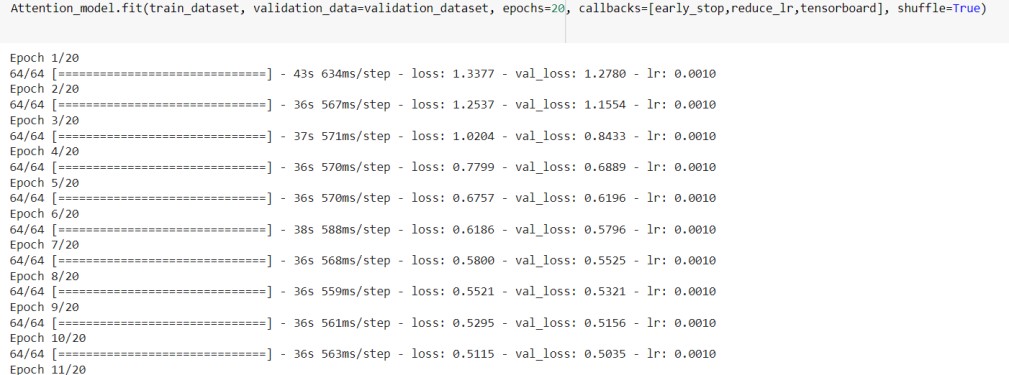
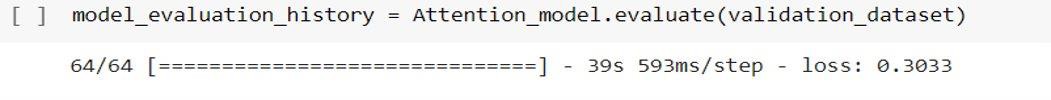


Fig 4.2 Training the model

#### B Training the model

The figure 4.1.2 depicts the Encoder Decoder Model with Attention is successfully trained with the input dataset and epoch takes time up to 600ms.After training the model we got an average loss rate of 0.0010 .



##### Fig 4.3 Evaluation of the model

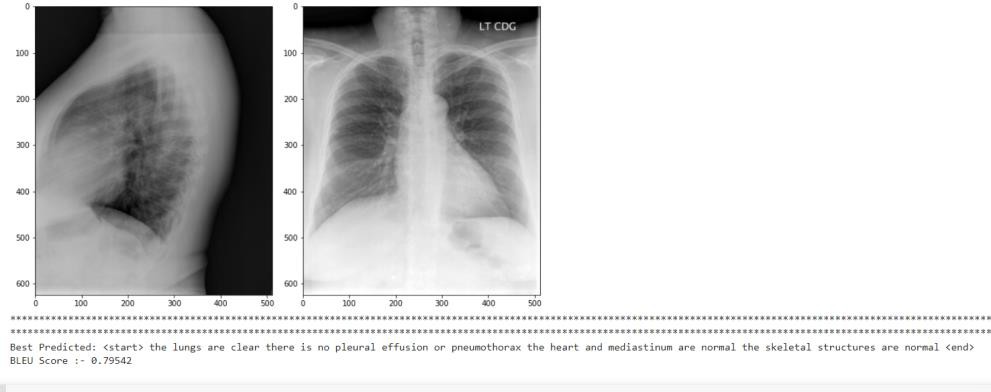
The figure 4.3 depicts that the Encoder Decoder Model with Attention is successfully evaluated with the validation dataset and got value loss of 0.3033

#### Test Case Results:



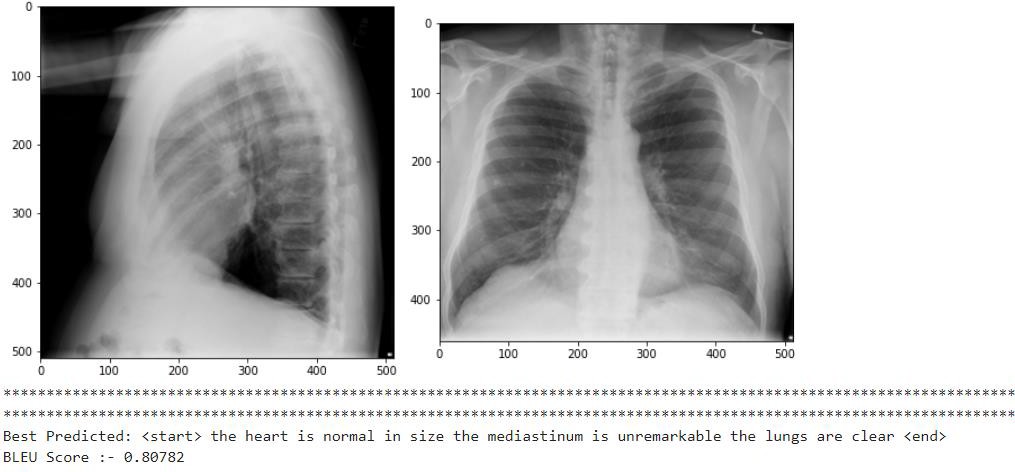
##### Fig 4.4 testcase 1

The figure 4.4 describes the report the heart size and pulmonary vascularity appear within normal limits the lungs are free of focal airspace disease no pleural effusion or pneumothorax is seen with an BLEU score of 0.79



##### Figure 4.5 test case 2

The figure 4.5 describes the report the lungs are clear there is no pleural effusion or pneumothorax the heart and mediastinum are normal the skeletal structures are normal with an BLEU score of 0.79



##### Fig:4.6 testcase 3

The figure 4.6 describes the report the heart size is normal in size the mediastinum is unremarkable the lungs are clear with an BLEU score of 0.80

* 1. **Observation from work**



##### Fig 4.7 no.of images in a report

The above shows the maximum number of images associated with a report and total images per count are

2-3208 ,1-446 ,3-181 ,0-104 ,4-15, 5-1 respectively

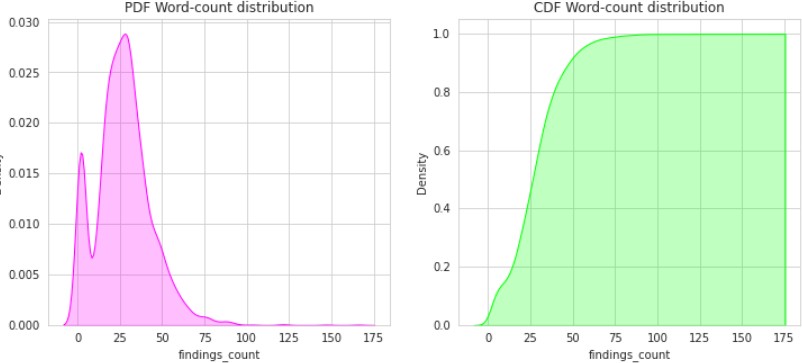


Figure 4.8 word count distribution

The above figure describes the pdf and cdf word count distribution

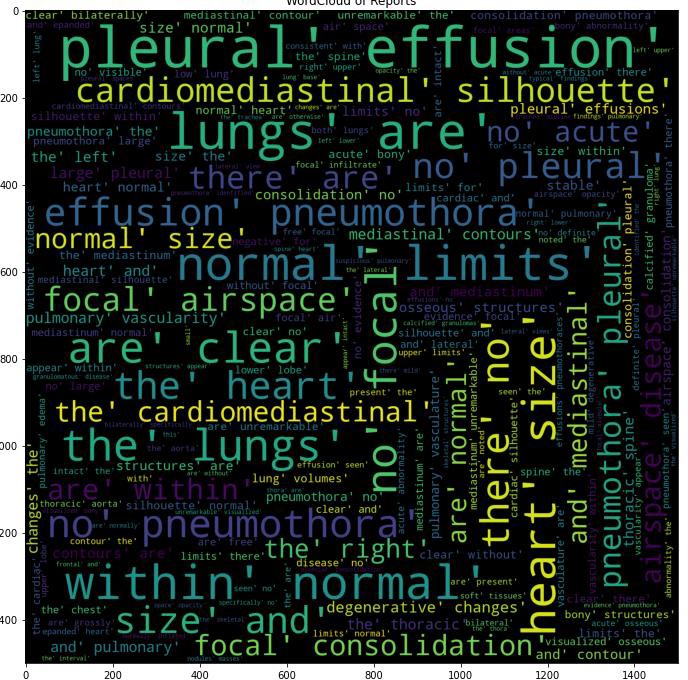
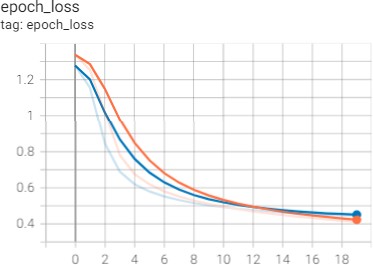


Fig 4.9 word cloud

The figure 4.9 describes the word cloud of text report after pre-processing which appears more number of times.



Fig;4.10 Train vs validation loss grap

The figure 4.10 depicts the loss graph of the Training and validation and validation loss a metric used to evaluate the deep learning model’s performance on the validation set.

* 1. **Conclusion**

### CHAPTER-5 CONCLUSION AND FUTURE WORK

The proposed model is an application to create automated textual reports for CXR, with the aim of assisting medical professionals in creating reports more efficiently and effectively. It is based on a LSTM feature extraction model that acts as an encoder that converts an image into a fixed act as an encoder that converts an image into a fixed- size vector representation, followed by an RNN decoder that generates corresponding sentences based on the learned image features. The effectiveness of the model was analyzed quantitatively and qualitatively on the CXR dataset.

#### Future Work

The future work of this project is the performance is also expected to increase when using a larger dataset by training on a greater number of images. No major hyperparameter tuning were done for any of the models. Therefore, a better

hyperparameter tuning might produce better results. Making use of little more advanced techniques like transformers or BERT, might yield better results.

#### References:

1. Sirshar M, Paracha MFK, Akram MU, Alghamdi NS, Zaidi SZY, Fatima T (2022) Attention based automated radiology report generation using CNN and LSTM.
2. Bustos A., Pertusa A., Salinas J. M., & de la Iglesia-Vayá M. (2020). Padchest: A large chest x-ray image dataset with multi-label annotated reports. Medical image analysis.
3. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., et al. (2019, July). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison.
4. Wang, X., Peng, Y., Lu, L., Lu, Z., & Summers, R. M. (2018). Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays.
5. Li, Y., Liang, X., Hu, Z., & Xing, E. P. (2018). Hybrid retrieval-generation reinforced agent for medical image report generation. In Advances in neural information processing systems.
6. . Zhang, Y. Xie, Z. Liao, G. Pang, J. Verjans,(2020) W. Li, Z. Sun, J. He, Y. Li, C. Shen, et al., Viral pneumonia screening on chest x-ray images using confidence-aware anomaly detection.
7. . Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U. Rajendra Acharya,(2019) Automated detection of covid-19 cases using deep neural networks with x-ray images.
8. .E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, arXiv preprint arXiv
9. Anderson, P., He, X., Buehler, C., Teney, D., Johnson, M., Gould, S., Zhang, L.(2018): Bottom-up and top-down attention for image captioning and visual question answering.
10. Jing, B., Xie, P., Xing, E(2018): On the automatic generation of medical imaging reports. arXiv preprint arXiv
11. Li, Y., Liang, X., Hu, Z., Xing, E.P (2018): Hybrid retrieval-generation reinforced agent for medical image report generation.
12. Vinyals, O., Toshev, A., Bengio, S., Erhan, D.: Show and tell: A neural image caption generator. In: Proceedings of the IEEE conference on computer vision and pattern recognition.

# DEPARTMENT OF INFORMATION TECHNOLOGY

**V.R.SIDDHARTHA ENGINEERING COLLEGE**

## PROJECT SUMMARY

|  |  |  |
| --- | --- | --- |
| **S.No** | **Item** | **Description** |
| **1** | **Project Title** | **Report generation on chest x-rays using deep learning** |
| **2** | **Student Names &**  **Numbers** | **M.J.N.V. SAI (208W1A12A0)**  **P. Monica (208W1A12A5)**  **V. Saiesh (208W1A12C7)** |
| **3** | **Name of The Guide** | **Dr. M Ashok Kumar** |
| **4** | **Name of The**  **Mentor** | **Dr. J Ebenezer** |
| **5** | **Research Group** | **Computer Vision and Remote Sensing** |
| **6** | **Application Area** | **Health Care** |
| **7** | **Aim of the Project** | **Automatic generation of a report on chest x-ray** |
| **8** | **Project Outcomes** | **The Goal of the project is to generate the text report on x-rays images which can reduce the time.** |

**Student Signatures**

* + - 1. **M.J.N.V. SAI**
      2. **P. Monica**
      3. **V. Saiesh**

**Signature of the guide:**