# **Sentiment Analysis of Twitter Feed Data Using CNN BERT and BiLSTM**

A PROJECT REPORT

#### Submitted by

**S S R SUBRAMANYA HEMANT KONDURI [RA2111050010010]**

**NAGARJUNA LEDALLA [RA2111050010019]**

#### Under the Guidance of

## **DR. M RAMPRASATH**

Assistant Professor, Department of Data Science and Business Systems

### *in partial fulfillment of the requirements* *for the degree of*

## BACHELOR OF TECHNOLOGY

## in

## COMPUTER SCIENCE ENGINEERING

## with specialization in BLOCKCHAIN TECHNOLOGY



## DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS

## COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

## KATTANKULATHUR- 603 203

### NOVEMBER 2024

Department of Computational Intelligence

##### SRM Institute of Science & Technology

##### Own Work\* Declaration Form

##### Degree/ Course :B.Tech in Computer Science and Engineering

**Student Name :S S R Subramanya Hemant Konduri, Nagarjuna Ledalla**

##### Registration Number :RA2111050010010, RA2111050010019

**Title of Work :Sentiment Analysis of Twitter Feed Data using CNN, BERT and BiLSTM**

I / We hereby certify that this assessment compiles with the University’s Rules and Regulations relating to Academic misconduct and plagiarism, as listed in the University Website, Regulations, and the Education Committee guidelines.

I / We confirm that all the work contained in this assessment is my / our own except where indicated, and that I / We have met the following conditions:

* Clearly referenced / listed all sources as appropriate
* Referenced and put in inverted commas all quoted text (from books, web, etc)
* Given the sources of all pictures, data etc. that are not my own
* Not made any use of the report(s) or essay(s) of any other student(s) either past or present
* Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)
* Compiled with any other plagiarism criteria specified in the Course handbook / University website

I understand that any false claim for this work will be penalized in accordance with the University policies and regulations.

|  |
| --- |
| **DECLARATION:** |
| I am aware of and understand the University’s policy on Academic misconduct and plagiarism and I certify that this assessment is my / our own work, except where indicated by referring, and that I have followed the good academic practices noted above.  S S R Subramanya Hemant Konduri[RA2111050010010]  Nagarjuna Ledalla[RA2111050010019]  Date: |
| S S R Subramanya Hemant Konduri[RA2111050010010]  Nagarjuna Ledalla[RA2111050010019] |



# SRM INSTITUTE OF SCIENCE AND TECHNOLOGY KATTANKULATHUR – 603 203

## BONAFIDE CERTIFICATE

Certified that 18CSP107L - Minor Project [18CSP108L- Internship] report titled “**PROJECT TITLE** ” is the bonafide work of “**STUDENT1 NAME [REG NUM], STUDENT2 NAME [REG NUM]”** who carried out the project work[internship] under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

**Dr.M.Ramprasath Dr. Kavitha V**

|  |  |  |
| --- | --- | --- |
| **SUPERVISOR**  **Assistant Professor**  DEPARTMENT OF  DATA SCIENCE AND BUSINESS SYSTEMS |  | **PROFESSOR &HEAD**  DEPARTMENT OF  DATA SCIENCE AND BUSINESS SYSTEMS |

**ACKNOWLEDGEMENTS**

We express our humble gratitude to **Dr. C. Muthamizhchelvan**, Vice-Chancellor, SRM InstituteofScience and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to **Dr. T. V. Gopal ,** Dean-CET, SRM Institute of Science and Technology, for his invaluable support.

We wish to thank **Dr. Revathi Venkataraman**, Professor and Chairperson, School of Com puting, SRM Institute of Science and Technology, for her support throughout the project work.

We encompass our sincere thanks to , **Dr. M. Pushpalatha**, Professor and Associate Chairperson, School of Computing and **Dr. C.Lakshmi,** Professor and Associate Chairperson, School of Computing, SRM Institute of Science and Technology, for their invaluable support.

We are incredibly grateful to our Head of the Department, **Dr. V. Kavitha**, Professor, Department of Computational Intelligence, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our Project Coordinators, ........., Panel Head, ......... and Panel Members, .........., Department of Computational Intelligence, SRM Institute of Science and Technology, for their inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor, Dr. Nadana Ravishankar, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to our guide, Dr. M. Ramprasath, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for providing us with an opportunity to pursue our project under his mentorship. He provided us with the freedom and support to explore the research topics of our interest. His / Her passion for solving problems and making a difference in the world has always been inspiring.

We sincerely thank all the staff and students of Department of Data Science and Business Systems, School of Computing, S.R.M Institute of Science and Technology, for their help during our project. Finally, we would like to thank our parents, family members, and friends for their unconditional love, constant support and encouragement

Authors

S S R Subramanya Hemant Konduri

Nagarjuna Ledalla

**ABSTRACT**

**TABLE OF CONTENTS**

[**ABSTRACT**](#_bookmark1) **I**

[**LIST OF FIGURES**](#_bookmark2) **IV**

**LIST OF TABLES V**

[**ABBREVIATIONS**](#_bookmark3) **VI**

**CHAPTER TITLE PAGE**

**NO. NO.**

[**1 INTRODUCTION**](#_bookmark4) **1**

* 1. General (Introduction to Project) 2
  2. Motivation 3
  3. Sustainable Development Goal of the Project 4

[**2 LITERATURE SURVEY**](#_bookmark20) **5**

2.1 subtitle 1 5

2.2 subtitle 2 10

2.3 Limitations Identified from Literature Survey (Research Gaps) 11

2.4 Research Objectives 12

2.5 Plan of Action (Project Road Map) 15

**3 RESEARCH METHODOLOGY 21**

3.3.3 Architecture Document 40

3.3.4 Outcome of objectives/ Result Analysis 41

42

**6 RESULTS AND DISCUSSIONS 40**

6.1 Project Outcomes (Performance Evaluation, Comparisons, Testing Results ) 41 6.2 Sub Title 42

**7 CONCLUSION AND FUTURE ENHANCEMENT 45**

**REFERENCES 46**

**APPENDIX**

**A CODING 48**

**B CONFERENCE PUBLICATION 50**

**C JOURNAL PUBLICATION 51**

**D PLAGIARISM REPORT 52**

**LIST OF FIGURES**

**CHAPTER TITLE PAGE**

**NO. NO.**

##### [2.1 Thresholding segmentation in action on the skin lesion image input](#_bookmark12) . . 4

[2.2 **Computer Vision Pipeline**](#_bookmark16). . . . . . . . . . . . . . . . . . . . . . . . . . 5

* 1. [**ROC curve CNN and dermatologists**](#_bookmark22) . . . . . . . . . . . . . . . . . . . . 8
  2. [**confusion matrix with CNN vs doctors**](#_bookmark26) . . . . . . . . . . . . . . . . . . . 10

[3.3 **Adam optimizer error rate**](#_bookmark29). . . . . . . . . . . . . . . . . . . . . . . . . 11

[4.1 **Sample lesions from each class**](#_bookmark32). . . . . . . . . . . . . . . . . . . . . . . 13

[4.2 **File structure of dataset**](#_bookmark34). . . . . . . . . . . . . . . . . . . . . . . . . . . 13

[4.3 **Diagnosis Techniques Plot**](#_bookmark36) . . . . . . . . . . . . . . . . . . . . . . . . . . 15

* 1. [**Number of data points of each class**](#_bookmark38). . . . . . . . . . . . . . . . . . . . 16

##### [Effect of SMOTE and number of variables on KNN (Euclidean Distance)](#_bookmark39) 17

[4.6 **YOLO CNN Layers**](#_bookmark43). . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

[4.7 **ANN sample architecture**](#_bookmark48). . . . . . . . . . . . . . . . . . . . . . . . . . 23

[4.8 **Convolution Layer simulation**](#_bookmark52). . . . . . . . . . . . . . . . . . . . . . . 25

[4.9 **Pooling layer simulation**](#_bookmark54) . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

[4.10 **CNN complete architecture**](#_bookmark56). . . . . . . . . . . . . . . . . . . . . . . . . 26

* 1. [**CNN pre assigned weights for each class**](#_bookmark58) . . . . . . . . . . . . . . . . . . 27
  2. [**Sample layers of our CNN model**](#_bookmark59). . . . . . . . . . . . . . . . . . . . . . 28

[5.1 **Overall architecture of Product.**](#_bookmark67). . . . . . . . . . . . . . . . . . . . . . 32

[5.2 **Difference between UI & UX**](#_bookmark69). . . . . . . . . . . . . . . . . . . . . . . . 33

[5.3 **Ionic Framework**](#_bookmark71). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34

[5.4 **Ionic Framework**](#_bookmark73). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

[5.5 **Back end Architecture**](#_bookmark75) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36

[5.6 **Components of API**](#_bookmark77). . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

[5.7 **REST API Architecture**](#_bookmark78). . . . . . . . . . . . . . . . . . . . . . . . . . . 38

[5.8 **SOAP API Architecture**](#_bookmark79). . . . . . . . . . . . . . . . . . . . . . . . . . . 39

vii

**LIST OF TABLES**

**CHAPTER TITLE PAGE**

**NO. NO.**

##### [2.1 Thresholding segmentation in action on the skin lesion image input](#_bookmark12) . . 4

[2.2 **Computer Vision Pipeline**](#_bookmark16). . . . . . . . . . . . . . . . . . . . . . . . . . 5

* 1. [**ROC curve CNN and dermatologists**](#_bookmark22) . . . . . . . . . . . . . . . . . . . . 8
  2. [**confusion matrix with CNN vs doctors**](#_bookmark26) . . . . . . . . . . . . . . . . . . . 10

**viii**

**ABBREVIATIONS**

**AES** Advanced Encryption Standard

**ANN** Artificial Neural Network **CNN** Colvonutional Neural Network **CSS** Cascading Style Sheet

**CV** Computer Vision

**DB** Database

**DNA** Deoxyribo Neucleic Acid

**GCP** Google Cloud Platform

**HAM** Human Against Machine **HTML** Hyper Text Markup Language **HTTP** Hyper Text Transfer Protocol **JS** Javascript

**KNN** K Nearest Neighbours

**MNIST** Modified National Institute of Standards and Technology

**PWA** Progressive Web App

**RNA** Ribo Neucleic Acid

**ROC** Receiver Operating Characteristic

**SASS** Syntactically Awesome Style Sheets **SMOTE** Synthetic Minority Oversampling Technique **SQL** Structured Query Language

**SVM** Support Vector Machine

**UI** User Interface

**UV** UltraViolet

**UX** User Experience

**YOLO** You Only Look Once

ix

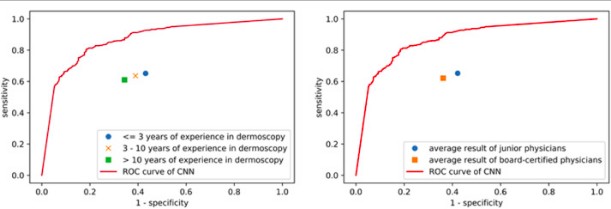
**CHAPTER 1**

**INTRODUCTION**

* 1. **Subtitle 1**

Font style for entire report must be times new roman. Chapter number and title must be capitalized with font size of 16pt bold. Subtitle 1 is 16pt with each word capitalized with bold. Subsection under a subtitle is 12pt with bold, each letter capitalized. All the content of the document is 12pt size with 1.5 Spacing. Left margin 1 inch and right margin 0.5 inch. Use justify option for both left and right alignment. A paragraph may contain maximum of 12 lines and an empty space to be left between each paragraph.

All figures must be numbered with corresponding chapter number like Fig 1.1, Fig 1.2 etc for chapter 1 and Fig 2.1, Fig 2.2 etc for chapter 2 and so on.. with figures in center alignment and the caption to be mentioned below the figure.



##### Fig 3.1: ROC curve CNN and dermatologists

All tables must be numbered with corresponding chapter number like Table 1.1, Table 1.2 etc for chapter 1 and Table 2.1, Table 2.2 etc for chapter 2 and so on.. with tables in center alignment and the caption to be mentioned above the Table.

Every chapter must start in new page. Page numbers are started from chapter 1 and to be included till the end of the report. Page numbers can be given at the bottom of the page with center alignment.

1

All references must be cited inside the text with sequential numbers[1] and to be listed in the same order in references. Two articles can be cited as [4,5] and multiple references can be cited as [7-10]. Minimum 15 references to be included.

Proofs must be included for all publications and plagiarism report to be generated using turnitin with the help of your guide with similarity index less than or equal to 10 percent. Total number of pages in the report is minimum 30 excluding coding and screenshots.

**REFERENCES**

1. Wang, Yili et al. (2022). Sentiment Analysis of Twitter Data.
2. Wankhade, K. et al. (2019). Sentiment Analysis on Twitter Data Using Machine Learning Techniques.
3. S. Symeonidis et al. (2018). Feature Engineering for Twitter Sentiment Analysis.
4. N. Jain et al. (2020). Deep Learning for Twitter Sentiment Analysis.
5. Bansal, S. et al. (2020). Sentiment Analysis on Twitter Data Using Bi-LSTM.
6. Silva, M. et al. (2020). Twitter Sentiment Analysis Using BERT.
7. Zhang, Y. et al. (2021). Attention Mechanisms in Sentiment Analysis for Tweets.
8. Patel, H. et al. (2020). Hybrid Machine Learning Approaches for Twitter Sentiment Analysis.
9. Gupte, M. et al. (2014). The Impact of Hashtags on Twitter Sentiment Analysis.
10. Go, A. et al. (2016). Twitter Sentiment Classification Using Distant Supervision.
11. Abdullah, M. et al. (2019). SVM and Neural Networks for Real-Time Twitter Sentiment Analysis.
12. Mittal, A. et al. (2019). Social Media Monitoring Using Sentiment Analysis of Tweets.
13. Agarwal, A. et al. (2011). Sentiment Analysis of Twitter Data Using Tree Kernels.
14. Dai, X. et al. (2019). Preprocessing Techniques for Twitter Sentiment Analysis.
15. Zubiaga, A. et al. (2018). Analyzing Sentiments in Real-Time Events via Twitter.
16. Wang, C. et al. (2016). Twitter Sentiment Analysis Using Recurrent Neural Networks.
17. Adwan, O. et al. (2018). Sentiment Analysis of Political Tweets.
18. Oliveira, A. et al. (2020). Twitter Sentiment Analysis for Brand Monitoring.
19. Mekala, R. et al. (2021). Transfer Learning for Twitter Sentiment Analysis.
20. Li, J. et al. (2019). Multi-Aspect Sentiment Analysis of Tweets Using CNN.
21. Asghar, M. et al. (2020). Comparative Analysis of ML Algorithms for TSA.
22. Cambria, E. et al. (2016). Affective Computing for Twitter Sentiment Analysis.
23. Liu, B. (2020). Lexicon-Based Sentiment Analysis in Twitter.
24. Hutto, C. et al. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text.
25. Yadav, V. et al. (2021). TSA Using Ensemble Learning.
26. Siddharth, N. et al. (2021). Long-Short Term Memory Networks for TSA.
27. Chowdhury, S. et al. (2019). Emotion Detection in Tweets Using BERT.
28. Habimana, T. et al. (2020). Twitter Sentiment Analysis for Disaster Response.

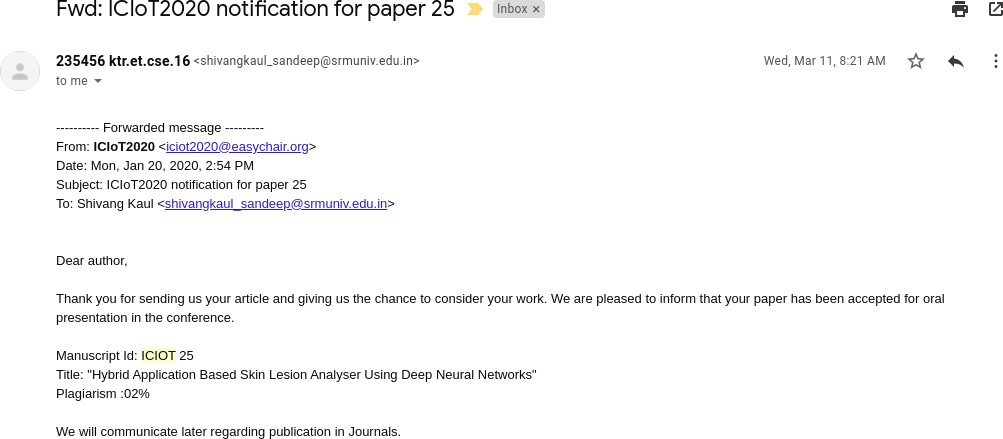
**APPENDIX A**

**CODING**

**APPENDIX B**

**CONFERENCE PRESENTATION**

Our paper on **Hybrid application based skin lesion analyzer using deep neural networks** was presented at ICIOT 2020 conference held at SRM. 200+ shortlisted teams presented their papers on various fields in the conference. Our paper got accepted as paper id : 25 with a plagiarism of just 2 %.



##### Figure A.1: ICIOT 2020 Acceptance

On presenting the paper in this international conference held at SRM KTR campus, we received positive remarks and suggestion from the judging panel. We were then awarded the best paper award at the same conference.



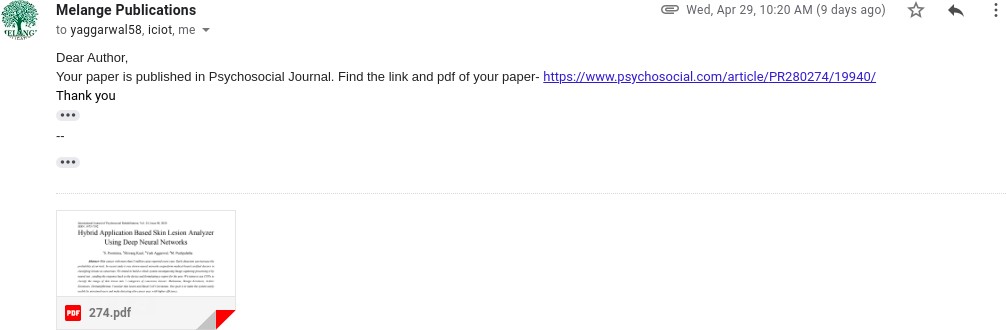
##### Figure A.2: ICIOT 2020 Best Paper award

83

**APPENDIX C**

**PUBLICATION DETAILS**

We submitted our research paper for publication at IJPR publication house puducherry. We had selected the journal **International Journal of Psychosocial Rehabilitation (ISSN: 1475- 7192)**. We got the acceptance notification from the IJPR stating our paper has been published in the April Issue of the same journal. Proof of publication is attached in figure [B.1](#_bookmark123) The research



##### Figure B.1: Publication Notification

paper cover page has been attached below.

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020 ISSN: 1475-7192

Hybrid Application Based Skin Lesion Analyzer Using Deep Neural Networks

1S. Poornima, 2Shivang Kaul, 3Yash Aggarwal, 4M. Pushpalatha

***Abstract--****Skin cancer with more than 5 million cases reported every year. Early detection can increase the probability of survival. In recent study it was shown neural networks outperform medical board certified doctors in classifying lesions as cancerous. We intend to build a whole system encompassing Image capturing processing it by neural net , sending the response back to the device and formulating a report for the user. We intent to use CNNs to classify the image of skin lesion into 7 categories of cancerous lesions: Melanoma, Benign Keratosis, Actinic Keratoses, Dermatofibroma, Vascular skin lesion and Basal Cell Carcinoma. Our goal is to make the system easily usable by untrained users and make detecting skin cancer easy with higher efficiency.*

***Key words--****Neural Networks, Image Processing, Convolu-tional Neural Networks, Skin Cancer Detection, Skin Lesion Imaging, App Development, Localization Algorithms, Cloud Computing, GCP, Compute Engine, App Engine.*

1. **INTRODUCTION**

Skin Cancer is a major kind of cancer with around 5 million reported cases worldwide every year. The major cause of skin cancer is exposure to UV rays. Diagnosing skin cancer generally included the skin lesion being examined by a doctor. Recent studies have shown neural networks to be more efficient in classifying lesion as cancerous as compared to trained doctors. Misdiagnosing or late detection of cancer can lead to a higher mortality rate and less chance of cure. The goal of this project is making detection and classification of lesions on the skin easier. Not all the marks on skin are a matter of concern but early detection and treatment of cancer can save lives. So this gives the user a way to check if there’s a chance of the mark on your skin being cancerous. The aim of this project is to detect and analyse such a correlation using neural networks. It is expected that the outcome of this project will lead to automated classification of skin lesions.

1. **LITERATURE SURVEY**

The following papers were read and analysed for the refer-ence of this paper. A brief image has been presented here.

1) Andre Esteva et al. 2017,” Dermatologist-level classification of skin cancer with deep neural networks.” Contribution: Claimed to classify skin lesions at par with board trained dermatologists. Methodology used:

*1Assistant Professor, CSE Department, SRMIST, Chennai, India*

*2Assistant Professor, CSE Department, SRMIST, Chennai, India,* [*shangkaul@gmail.com*](mailto:shangkaul@gmail.com) *3Assistant Professor, CSE Department, SRMIST, Chennai, India,* [*yaggarwal58@gmail.com*](mailto:yaggarwal58@gmail.com) *4Assistant Professor, CSE Department, SRMIST, Chennai, India*

**DOI: 10.37200/IJPR/V24I8/PR280274**

**Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020 2545**

**APPENDIX D**

**PLAGIARISM REPORT**