



**FINAL YEAR PROJECT REPORT**

# **CLAIM CRACKER – A CLAIM DETECTOR AND VALIDATOR**

**In fulfillment of the requirement  
For degree of  
BS (COMPUTER SCIENCES)**

**By**

|                                |                     |
|--------------------------------|---------------------|
| <b>MUHAMMAD SHAHMEER NASIR</b> | <b>67748 (BSCS)</b> |
| <b>HARIS MUSA</b>              | <b>67714 (BSCS)</b> |
| <b>SHEPHER ABRAHAM GILL</b>    | <b>65039 (BSCS)</b> |

**SUPERVISED**

**BY**

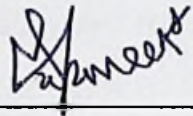
**MS. SAMEENA JAVAID**

**BAHRIA UNIVERSITY (KARACHI CAMPUS)**

**FALL-2023**

**DECLARATION**

We hereby declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Signature : 

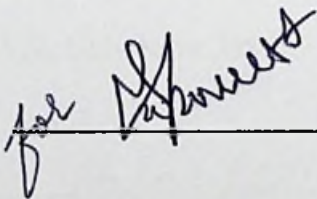
Name : Muhammad Shahmeer Nasir

Reg No. : 67748

Signature : Haris

Name : Haris Musa

Reg No. : 67714

Signature : 

Name : Shepherd Abraham Gill

Reg No. : 65039

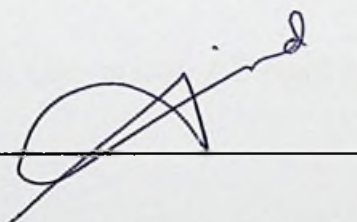
Date : 16/01/24

### APPROVAL FOR SUBMISSION

We certify that this project report entitled **“CLAIM CRACKER – A CLAIM DETECTOR AND VALIDATOR”** was prepared by **MUHAMMAD SHAHMEER NASIR, HARIS MUSA and SHEPHER ABRAHAM GILL** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Science in **Computer Science** at Bahria University, Karachi.

Approved by,

Signature :



Supervisor : Ms. Sameena Javaid

Date

: 16/01/24.

The copyright of this report belongs to Bahria University according to the Intellectual Property Policy of Bahria University BUORIC-P15 amended on April 2019. Due acknowledgement shall always be made of the use of any material contained in, or derived from, this report.

© 2023 Bahria University. All right reserved.

## **ACKNOWLEDGEMENT**

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express our gratitude to our project supervisor, Ms. Sameena Javaid for her invaluable advice, guidance and her enormous patience throughout the development of the project.

In addition, we would also like to express our gratitude to our loving parents and friends who had helped and given us encouragement.

## **CLAIM CRACKER – A CLAIM DETECTOR AND VALIDATOR**

### **ABSTRACT**

The modern digital era produces a large volume of data on an hourly basis around the world including Pakistan. This data contains a significant amount of false, misleading, and fake information and claims. One of the major reasons for the unstable political situation in Pakistan is the spreading of fake claims. These fake claims result in a quick fire around the public causing unfavourable events. Although these fake claims may get exposed as false or invalid claims later, the damage had been done till then. The reason is the manual research for the validation of these claims which takes lot of time to find out the truth. Automated claim detectors and validators can fill this gap. It will result in preventing serious damages by providing the reality of claims in a matter of few minutes. By using advanced natural language processing and machine learning techniques which involve Recurrent Neural Network (RNN), LSTM, Bi-LSTM, CNN and GRU, we have prepared a model trained on past Pakistani political and influential claims and their reality. It is capable of detecting claims in sentences and validating them. To make this model easily accessible to everyone we have made a ReactJS-based modern web application with the Django backend along with the ClaimBuster, arXiv and OpenAI API. It integrates the ClaimCracker model with the interactive user interface of the ClaimCracker web application. This dynamic web application provides a way for the end user to detect and validate claims with just a single button click. The model is capable of providing 93% accuracy for detecting and validating claims based on Pakistani politics and current affairs.

## TABLE OF CONTENTS

|                                        |              |
|----------------------------------------|--------------|
| <b>DECLARATION</b>                     | <b>ii</b>    |
| <b>APPROVAL FOR SUBMISSION</b>         | <b>iii</b>   |
| <b>ACKNOWLEDGEMENT</b>                 | <b>v</b>     |
| <b>ABSTRACT</b>                        | <b>vi</b>    |
| <b>TABLE OF CONTENTS</b>               | <b>vii</b>   |
| <b>LIST OF TABLES</b>                  | <b>xi</b>    |
| <b>LIST OF FIGURES</b>                 | <b>xii</b>   |
| <b>LIST OF SYMBOLS / ABBREVIATIONS</b> | <b>xviii</b> |
| <b>LIST OF APPENDICES</b>              | <b>xix</b>   |

## CHAPTER

|          |                                   |          |
|----------|-----------------------------------|----------|
| <b>1</b> | <b>INTRODUCTION</b>               | <b>1</b> |
|          | 1.1 Background                    | 1        |
|          | 1.2 Problem Statement             | 2        |
|          | 1.3 Aims and Objectives           | 2        |
|          | 1.4 Scope of Project              | 2        |
| <b>2</b> | <b>LITERATURE REVIEW</b>          | <b>4</b> |
|          | 2.1 Deep Learning                 | 4        |
|          | 2.2 ClaimBuster                   | 7        |
|          | 2.3 React                         | 7        |
|          | 2.4 Django                        | 8        |
| <b>3</b> | <b>REQUIREMENT SPECIFICATIONS</b> | <b>9</b> |
|          | 3.1 NLP Model                     | 9        |

|          |                                                      |           |
|----------|------------------------------------------------------|-----------|
| 3.2      | Deep Learning Techniques                             | 9         |
| 3.3      | Web Application (Frontend – ReactJS)                 | 9         |
| 3.4      | Backend - Django                                     | 10        |
| 3.5      | Integration with ClaimBuster’s API                   | 10        |
| 3.6      | Accuracy and Performance Metrics                     | 10        |
| 3.7      | Documentation                                        | 11        |
| 3.8      | Testing                                              | 11        |
| 3.9      | Scalability and Future Development                   | 11        |
| 3.10     | Security                                             | 11        |
| <b>4</b> | <b>DESIGN AND METHODOLOGY</b>                        | <b>12</b> |
| 4.1      | Product Development Life Cycle (PDLC)                | 12        |
| 4.2      | Technological Priorities                             | 14        |
| 4.3      | Technology Stack                                     | 14        |
| 4.4      | Frontend – React                                     | 14        |
| 4.5      | Claim Validation System                              | 15        |
| 4.5.1    | Data Pre-processing                                  | 15        |
| 4.5.2    | Embedding Layer                                      | 15        |
| 4.5.3    | Deep Learning Architectures                          | 16        |
| 4.5.4    | Combined Approach                                    | 22        |
| 4.5.5    | Evaluation Metrics                                   | 22        |
| 4.6      | ClaimBuster API                                      | 24        |
| 4.7      | Natural Language Processing (NLP) and Neural Network | 24        |
| <b>5</b> | <b>SYSTEM IMPLEMENTATION</b>                         | <b>26</b> |
| 5.1      | Data Scraping                                        | 26        |
| 5.2      | Technology Stack                                     | 28        |
| 5.3      | Model Training                                       | 29        |
| 5.3.1    | Data Pre-processing and Word Cloud Visualization     | 29        |
| 5.3.2    | Preparing the Data for Model Input                   | 29        |
| 5.3.3    | Word2Vec for Word Embeddings                         | 29        |
| 5.3.4    | Deep Learning Models                                 | 29        |

|          |                                                 |           |
|----------|-------------------------------------------------|-----------|
| 5.3.5    | Hyperparameters and Training                    | 30        |
| 5.3.6    | Ensemble Model                                  | 34        |
| 5.3.7    | Model Evaluation                                | 35        |
| 5.3.8    | Visualization of Model Performance              | 35        |
| 5.4      | APIs                                            | 35        |
| 5.4.1    | ClaimBuster API                                 | 35        |
| 5.4.2    | Arxiv API for Deep Learning Paper Fetching      | 35        |
| 5.4.3    | OpenAI API for Predictive Question Answering    | 36        |
| 5.4.4    | Predict API Endpoint Using Keras and TensorFlow | 36        |
| 5.5      | ReactJS Frontend Development                    | 36        |
| 5.5.1    | Tailwind CSS                                    | 37        |
| <b>6</b> | <b>SYSTEM TESTING AND EVALUATION</b>            | <b>38</b> |
| 6.1      | Dataset Description                             | 38        |
| 6.2      | Data Analysis                                   | 38        |
| 6.3      | Evaluation Metrics                              | 40        |
| 6.3.1    | LSTM Evaluation                                 | 40        |
| 6.3.2    | Bi-LSTM Evaluation                              | 43        |
| 6.3.3    | RNN Evaluation                                  | 46        |
| 6.3.4    | CNN Evaluation                                  | 49        |
| 6.3.5    | GRU Evaluation                                  | 52        |
| 6.3.6    | Ensemble Model Evaluation                       | 55        |
| 6.4      | Machine Learning Algorithms' Accuracy           | 57        |
| 6.5      | Results of Models                               | 58        |
| 6.5.1    | LSTM Model                                      | 58        |
| 6.5.2    | Bi-LSTM Model                                   | 59        |
| 6.5.3    | RNN Model                                       | 59        |
| 6.5.4    | GRU Model                                       | 60        |
| 6.5.5    | Ensemble Model                                  | 60        |
| 6.6      | Frontend - React                                | 61        |
| 6.6.1    | API and Backend Testing with Postman            | 63        |