



**PhD in Information Technology and Electrical Engineering**  
Università degli Studi di Napoli Federico II

**PhD Student: AREEBA UMAIR**

---

**Cycle: XXXVI**

**Training and Research Activities Report**

**Year: First**

*Areeba Umair*

**Tutor: prof. Elio Masciari**

*Elio Masciari*

**Co-Tutor: Prof. Vincenzo Moscato**

**Date: October 28, 2021**

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

## 1. Information:

- **PhD student:** Areeba Umair
- **DR number:** DR995146
- **Date of birth:** 07/01/1996
- **Master Science degree:** MSCS                      **University:** NUCES, Pakistan
- **Doctoral Cycle:** XXXVI
- **Scholarship type:** UNINA
- **Tutor:** Prof. Elio Masciari
- **Co-tutor:** Prof. Vincenzo Moscato

## 2. Study and training activities:

| Activity   | Type <sup>1</sup> | Hours | Credits | Dates                              | Organizer   | Certificate <sup>2</sup> |
|--|-------------------|-------|---------|------------------------------------|---|--------------------------|
| Digital Forensics; methods, practices and tools                | Ad hoc Course     | 10    | 3       | 03-05-06-09-10 / 11/2020           | Dr. Giovanni Cozzolino                                  | Y                        |
| Statistical data analysis for science and engineering research | Ad Hoc Course     | 12    | 4       | 17-19-24-25 / 02 / 03-04 / 03/2021 | Prof. Roberto Pietrantuono                              | Y                        |
| Hardware and Software Architectures for Big Data – Mod. B      | Msc. Course       |       | 6       | 18/06/2021                         | Prof. Vincenzo Moscato                                  | Y                        |
| Big Data Analytics and Business Intelligence                   | Msc. Course       |       | 6       | 22/06/2021                         | Prof. Giancarlo Sperli'                                 | Y                        |
| 2021 Spring School on Transferable Skills                      | Doctoral School   |       | 2       | 4-5 May 2021                       | Department of Pharmacy University of Naples Federico II | Y                        |
| AIRO PhD School 2021 and 5th AIRO Young Workshop               | Doctoral School   |       | 3.6     | 8-12 February 2021                 | Sterle C., Sforza A., Boccia M., Masone A.              | Y                        |
| Safety Training Course   | Course            | 16    | 0       | May 2021                           | Dott.ssa Liliana Lista                                  | Y                        |
| Robot  | Seminar           | 2.5   | 0.5     | 17                                 | Prof. Paolo   | Y                        |

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

|   |         |     |     |                  |                                   |   |
|---|---------|-----|-----|------------------|-----------------------------------|---|
| Manipulation and Control  |         |     |     | November 2020    | Dario, Scuola Sant'Anna Pisa      |   |
| Patent Searching Best Practices with IEEE Xplore  | Seminar | 1   | 0.2 | 27 November 2020 | Rachel Berrington                 | Y |
| How to get published with the IEEE  | Seminar | 1.5 | 0.3 | 02 December 2020 | Rachel Berrington                 | Y |
| Advances in Machine Learning for Modelling and Understanding in Earth Sciences  | Seminar | 1.5 | 0.3 | 27 January 2021  | Prof. Antonio Iodice              | Y |
| NLP ED AI NEL MONDO ENTERPRISE  | Seminar | 2   | 0.4 | 21 January 2021  | Prof. Marco PASSAROTTI            | Y |
| AI: Artificial Intelligence for notary's sector - a case study  | Seminar | 1   | 0.2 | 27 January 2021  | Salvatore Palange                 | Y |
| Machine Learning: Causality loss in translation   | Seminar | 1.5 | 0.3 | 10 February 2021 | Edwin A. Valentijn                | Y |
| Approaches to graph machine learning  | Seminar | 1.5 | 0.3 | 17 February 2021 | Miroslav Cepek                    | Y |
| IEEE Authorship and Open Access Symposium: Best Practices to Get Published to Increase the Exposure and Impact of Your Research | Seminar | 1.5 | 0.3 | 22 April 2021    | Rachel Berrington                 | Y |
| Wireless Intelligence: From Reconfigurable Surfaces to Edge/Cloud Communications  | Seminar | 4   | 0.8 | 26 March 2021    | Prof. M. Martalò, Prof. R. Raheli | Y |
| Logic-based Learning of   | Seminar | 1   | 0.2 | 08 April 2021    | Prof. Mark Law                    | Y |

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

|   |         |     |     |               |                                       |   |
|---|---------|-----|-----|---------------|---------------------------------------|---|
| Answer Set Programs   |         |     |     |               |                                       |   |
| Why Do We Cooperate? Understanding and Modelling Societies using Reinforcement Learning               | Seminar | 1.5 | 0.3 | 01 April 2021 | Prof. Mirco Musolesi                  | Y |
| Visual Interaction and Communication in Data Science  | Seminar | 2   | 0.4 | 03 March 2021 | Prof. Marco Quartulli                 | Y |
| Big Data and Computational Linguistics  | Seminar | 2   | 0.4 | 10 March 2021 | Prof. Francesco Cutugno               | Y |
| SENSORIA Health   | Seminar | 1   | 0.2 | 17 March 2021 | Prof. Stefano Rossotti                | Y |
| The Science of Science  | Seminar | 1.5 | 0.3 | 25 March 2021 | Prof. Giuseppe Longo                  | Y |
| Distributional Semantics Methods: How Linguistic features can improve the semantic representation     | Seminar | 1.5 | 0.3 | 28 April 2021 | Prof. Alessandro Maisto               | Y |
| Risk assessment in real life: experiences from the railway domain                                     | Seminar | 1.5 | 0.3 | 26 May 2021   | Prof. Valeria Vittorini               | Y |
| Synchronization in complex networks, hypergraphs and simplicial complexes                             | Seminar | 1   | 0.2 | 27 May 2021   | Dr. Marco Coraggio, Dr. Micol Benetti | Y |
| Optimized Graph Representations for Right-Wing Reddit Community Detection Using Graph Neural Networks | Seminar | 1   | 0.2 | 30 April 2021 | Prof.ssa Silvia Rossi.                | Y |
| INTRODUCTION TO LEGGED ROBOTS AND   | Seminar | 2   | 0.4 | 26 May 2021   | Dr. Fabio Ruggiero                    | Y |

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

|  |          |     |     |              |                       |   |
|--|----------|-----|-----|--------------|-----------------------|---|
| EXAMPLES OF IIT'S DYNAMIC LEGGED SYSTEMS LA  |          |     |     |              |                       |   |
| INTRODUCTION TO UNDERWATER ROBOTICS  | Seminar  | 2   | 0.4 | 18 May 2021  | Dr. Fabio Ruggiero    | Y |
| Ethics of Quantification   | Seminar  | 2   | 0.4 | 26 May 2021  | Andrea Saltelli       | Y |
| Sadas Engine, an innovative DBMS for the DATA WAREHOUSE, great performance in the VLDB environment.  | Seminar  | 1.5 | 0.3 | 23 June 2021 | Luca De Rosa          | Y |
| 5G: Esposizione ai Campi Elettromagnetici e Metodologie di Misura  | Seminar  | 4   | 0.8 | 16 July 2021 | Prof. Nicola Pasquino | Y |
| Study on Fake News detection<br>Submission of article in Applied Sciences Journal<br>Addressing the comments of reviewers for the journal paper.<br>Publishing the final version of the paper. | Research |     | 10  |              |                       |   |
| Study on topic modelling<br>Study on transformers and attention mechanism<br>Completed review of about 30 summaries over transformers and attention mechanism                                  | Research |     | 5   |              |                       |   |
| Study on transformers and attention mechanisms.<br>Start writing survey paper for sentimental analysis during COVID-   | Research |     | 5   |              |                       |   |

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

|   |          |  |   |  |  |  |
|---|----------|--|---|--|--|--|
| 19.   |          |  |   |  |  |  |
| Submitted survey paper for sentimental analysis during COVID-19 in Ideas 2021 Conference. | Research |  | 5 |  |  |  |
| Acceptance of conference paper  | Research |  | 5 |  |  |  |
| Submission of paper in Journal  | Research |  | 5 |  |  |  |

1) Courses, Seminar, Doctoral School, Research, Tutorship

2) Choose: Y or N

## 2.1. Study and training activities - credits earned

|                 | Courses     | Seminars   | Research  | Tutorship | Total       |
|-----------------|-------------|------------|-----------|-----------|-------------|
| Bimonth 1       | 3           | 1          | 10        | 0         | 14          |
| Bimonth 2       | 3.6         | 1.5        | 5         | 0         | 10.1        |
| Bimonth 3       | 4           | 3.2        | 5         | 0         | 12.2        |
| Bimonth 4       | 14          | 2.2        | 5         | 0         | 21.2        |
| Bimonth 5       | 0           | 0.8        | 5         | 0         | 5.8         |
| Bimonth 6       | 0           | 0          | 5         | 0         | 5           |
| <b>Total</b>    | <b>24.6</b> | <b>8.7</b> | <b>35</b> | <b>0</b>  | <b>63.3</b> |
| <b>Expected</b> | <b>20</b>   | <b>5</b>   | <b>35</b> | <b>0</b>  |             |

## 3. Research activity:

### Research Topic:

My research topic is “Devising Artificial Intelligence Tools for Complex Data”. During my first year of Ph.D., I worked on the COVID-19 sentimental analysis. I performed two research activities in the first year. Sentimental analysis is the emerging field in text mining where people’s feeling, and emotions are extracted using different techniques [2]. COVID-19 has declared as pandemic and effected people’s lives all over the globe. It caused the feelings of fear, anxiety, anger, depression, and many other psychological issues.

### 3.1. Sentimental Analysis Applications and Approaches during COVID-19: A Survey

I began my research by reviewing the current literature on COVID sentimental analysis. For that purpose, the sentimental analysis applications and methods which are used for COVID-19 research were explored. I have performed the survey of thirty primary studies related to sentimental analysis during COVID-19 pandemic and figure out the techniques that have been applied in order to classify the sentiments of the people as well as the application areas of sentimental analysis during COVID research. The objectives of the survey were:

- to identify the data sources and data volume of sentimental analysis during COVID-19,

- to identify the mostly used approaches and
- to identify the mostly used applications of sentimental analysis during COVID.
- Future implications of research with respect to COVID

The statistics shows that 24 out of 30 studies uses twitter as a data source while other sources of data are online media and forums, Weibo account, WeChat account, Reddit, Yelp, RateMDs, HealthGrades, and Vitals and Qingbo Big Data Agency. Twitter is considered as most popular social media platform having almost 81.47 million registered users [1]. Sentimental Classification Approaches, during COVID-19 research, can be divided into three types. Machine learning based approaches, lexicon-based approaches, and hybrid approaches. The comparison of thirty primary studies shows that Naive Bayes and SVM are the widely used algorithms of sentimental analysis for COVID-19 research. The applications of sentimental analysis during COVID includes the analysis of people's sentiments specially students, reopening sentiments, analysis of restaurants reviews and analysis of vaccine sentiments.

The survey concludes that the sentimental analysis during COVID-19 is still an open field and contains many interesting topics using advanced methods of machine learning and deep learning.

### 3.2. Sentimental Analysis of tweets towards COVID-19 Vaccines Using BERT model

Vaccine hesitancy is the main issue after COVID-19 pandemic and is defined as a delay in acceptance or refusal of vaccine despite its availability by the WHO Strategic Advisory Group of Experts on Immunization. The vaccine hesitancy and misinformation are hurdle in many countries in the control and prevention of many diseases. In the case of COVID-19, anti-vaccine activist is campaigning to demotivate the need of vaccines while some are denying the existence of COVID-19 till now [3]. To ensure the acceptance of COVID vaccine among the population, the government should understand the people sentiments towards vaccine [4].

#### **Proposed Framework:**

Proposed framework consists of four phases:

- Phase one deals with the dataset collection and preo-processing. Dataset gathered from tweets is usually in raw form. There is always a need to convert the data into more useful form. Hence, data pre-processing is performed in order to achieve the good accuracy.
- Phase two is social network analysis. It focuses on the world in terms of different overlying networks that connect for friendship, information, money and power. It leads to new realization about culture, history, politics and other topics. Connections of people have a huge factor over their way of life and their career choices. Social network analysis is more realistic because it can establish a relationship between the people and connections at the same time.
- Phase three includes the extraction of sentiments from the text and finding the semantic polarity of the extracted sentiments leading to the sentimental classification.
- Phase four is sentimental classification using BERT model.

# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

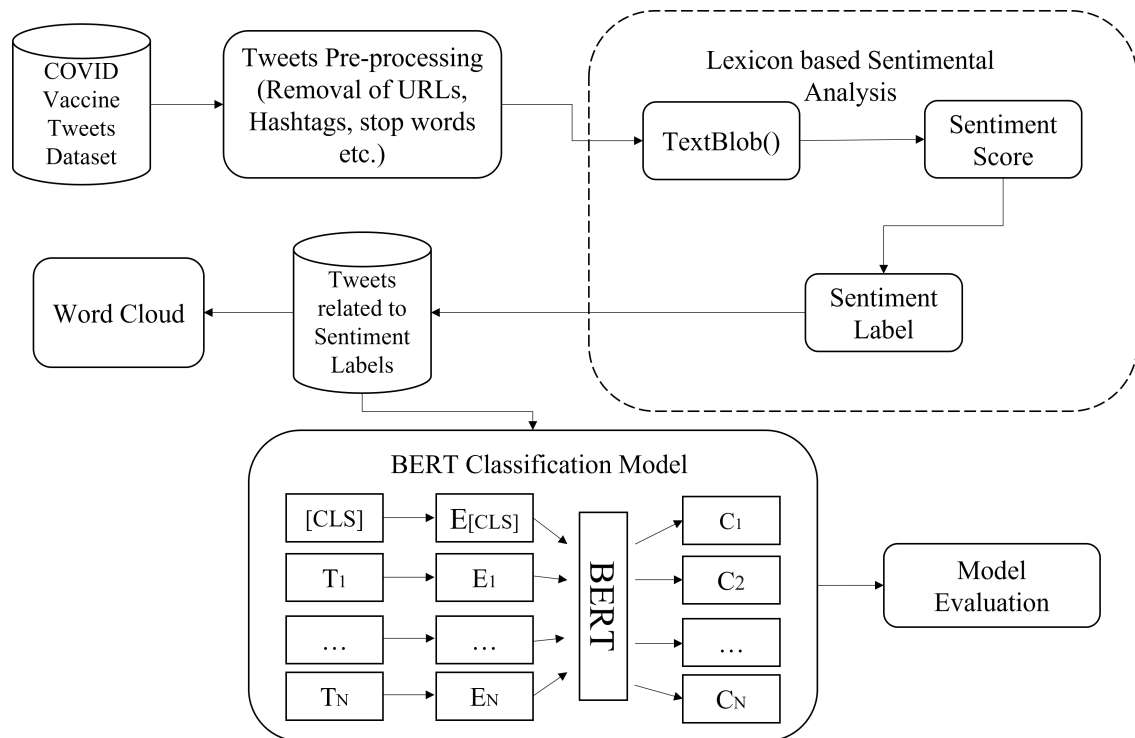


Figure 1 Proposed Framework

## Results:

The results shows that our proposed BERT model outperformed the state-of-the-art machine learning model for positive as well as negative sentiment classification by achieving maximum precision, recall and F-measure. BERT achieved 55 precision, 69 recall and 58 F-score in the case of positive tweet classification while it achieved 54 precision, 85 recall and 64 F-score for negative tweet classification. These results are highest amongst all other state-of-the-art algorithms.

## 4. Research products:

1. Umair, Areeba, et al. "Spatiotemporal Analysis of Web News Archives for Crime Prediction." *Applied Sciences* 10.22 (2020): 8220. (Accepted, IF= 2.679)
2. Umair, A., Masciari, E., & Habib Ullah, M. H. (2021, July). *Sentimental Analysis Applications and Approaches during COVID-19: A Survey*. In *25th International Database Engineering & Applications Symposium* (pp. 304-308). (Accepted B Ranked conference)
3. Areeba Umair, Elio Masciari, "Using High Performance Approaches to Covid-19 Vaccines Sentiment Analysis" *30th Euromicro International Conference on Parallel, Distributed and Network-based Processing 2021* (Submitted)
4. Areeba Umair, Elio Masciari, "Artificial Intelligence Based Analysis of Positive and Negative Tweets Towards COVID-19 Vaccines" *IEEE International Conference on Bioinformatics & Biomedicine, 2021* (Submitted)



# Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI

Author: Areeba Umair

---

## 5. Conferences and seminars attended

25<sup>th</sup> International Database Engineering and Applications Symposium, IDEAS, Montreal, Canada, Jul 14, 2021 - Jul 16, 2021. (**Presented my paper**)

## 6. Activity abroad:

*During my first year, I did not spend my time abroad.*

## 7. Tutorship

*During my first year, I did not perform tutorship.*

## References:

- [1] Hassan Adamu, Syaheerah Lebai Lutfi, Nurul Hashimah Ahamed Hassain Malim, Rohail Hassan, Assunta Di Vaio, and Ahmad Sufiril Azlan Mohamed. 2021. Framing twitter public sentiment on Nigerian government COVID-19 palliatives distribution using machine learning. Sustain. 13, 6 (2021). <https://doi.org/10.3390/su13063497>
- [2] Giuseppe Manco, Elio Masciari, and Andrea Tagarelli. 2002. A Framework for Adaptive Mail Classification. In 14th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2002), 4-6 November 2002, Washington, DC, USA. IEEE Computer Society, 387. <https://doi.org/10.1109/TAI.2002.1180829>
- [3] Jerrey V. Lazarus et al. \A global survey of potential acceptance of a COVID-19 vaccine". In: Nat. Med. 27.2 (2021), pp. 225{228. issn: 1546170X. doi: 10.1038/s41591-020-1124-9. url: <http://dx.doi.org/10.1038/s41591-020-1124-9>.
- [4] Holly Seale et al. \Examining Australian public perceptions and behaviors towards a future COVID-19 vaccine". In: medRxiv (2020), pp. 1{9. issn:1471-2334. doi: 10.1101/2020.09.29.20204396.