



# **DEEP LEARNING FOR ARABIC TEXT SENTIMENT ANALYSIS**

By

**Rana Mahmoud Kamel AbdelMoneim Kamel**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communication Engineering**

FACULTY OF ENGINEERING, CAIRO UNIVERSITY  
GIZA, EGYPT  
2019

# **DEEP LEARNING FOR ARABIC TEXT SENTIMENT ANALYSIS**

By

**Rana Mahmoud Kamel AbdelMoneim Kamel**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communication Engineering**

Under the Supervision of

**Prof. Dr. Mohamed Nafie**

Professor

Electronics and Communication Engineering Department

Faculty of Engineering, Cairo University

**Dr. Karim Seddik**

Associate Professor

Communications and Computer department

Faculty of Engineering, American University in Cairo

**FACULTY OF ENGINEERING, CAIRO UNIVERSITY**

**GIZA, EGYPT**

**2019**

# **DEEP LEARNING FOR ARABIC TEXT SENTIMENT ANALYSIS**

By

**Rana Mahmoud Kamel AbdelMoneim Kamel**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communication Engineering**

Approved by the Examining Committee:

---

**Prof. Dr. Mohamed Nafie**, Thesis Main Advisor

---

**Dr. Karim Seddik**, Advisor

Associate Professor at faculty of engineering, American University in Cairo

---

**Prof. Dr. Mohamed Khairy**, Internal Examiner

---

**Prof. Dr. Samhaa El-Beltagy**, External Examiner

Professor at faculty of Computer Science, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY

GIZA, EGYPT

2019

**Engineer's Name:** Rana Mahmoud Kamel AbdelMoneim Kamel  
**Date of Birth:** 01/03/1991  
**Nationality:** Egyptian  
**E-mail:** rana.mahmoudkamel@hotmail.com  
**Phone:** +20 100 387 835 3  
**Address:** 108, St. 104, Maadi, Cairo  
**Registration Date:** 01/03/2015  
**Awarding Date:** 2019  
**Degree:** Master of Science  
**Department:** Electronics and Communication Engineering



**Supervisors:**

Prof. Dr. Mohamed Nafie  
Dr. Karim Seddik

**Examiners:**

Prof. Dr. Mohamed Nafie	(Thesis main advisor)
Dr. Karim Seddik	(Advisor)
Associate Professor at faculty of Engineering, American University in Cairo	
Prof. Dr. Mohamed Khairy	(Internal examiner)
Prof. Dr. Samhaa El-Beltagy	(External examiner)
Professor at faculty of Computer Science, Cairo University	

**Title of Thesis:**

Deep Learning for Arabic Text Sentiment Analysis

**Key Words:**

Deep Learning; Word Embeddings; Arabic Natural Language processing;  
Sentiment Analysis

**Summary:**

Nowadays, people express their opinions, reviews on products, movies, hotels, restaurants... etc publicly on the internet on social media platforms, blogs or forums. The number of Arabic speaking users on the internet has been increased in the last decade and the research for analyzing the Arabic text has gained a lot of attention.

In my work, deep learning techniques are used to classify Arabic tweets and reviews into two classes (positive/ negative) and three classes (positive/ negative/ neutral). Also, this work investigates whether deep learning can overcome ordinary machine learning algorithms and replace the effort of feature engineering in previous work. Finally, deep learning proved to have better results than machine learning techniques for most of the used datasets by using a data augmentation architecture.

# Acknowledgements

This thesis would not exist without the support of many people. First of all, I would like to thank my supervisors Prof. Dr. Mohamed Nafie and Prof. Karim Seddik for the support provided during my work and reviewing the thesis.

I would like to dedicate a special thanks to Prof. Dr. Samhaa EL-beltagy for the continuous support and help I received during my thesis development as well as the encouragement and helpful feedback.

Finally, I would like to thank my lovely parents and husband for the continuous support and encouragement during my studies and specially in the hard times I passed by. I couldn't have done it without you.

# Disclaimer

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute.

I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Rana Mahmoud Kamel AbdelMoneim Kamel

Date: 10/6/2019

Signature:

# Table of Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Disclaimer</b>	<b>ii</b>
<b>Table of Contents</b>	<b>iii</b>
<b>List of Tables</b>	<b>v</b>
<b>List of Figures</b>	<b>vi</b>
<b>Abstract</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Problem Formulation . . . . .	1
1.2 Motivation . . . . .	2
1.3 Thesis Contribution . . . . .	2
1.4 Thesis Organization . . . . .	3
<b>2 Theoretical Background</b>	<b>4</b>
2.1 Natural Language Processing . . . . .	4
2.1.1 NLP and machine Learning . . . . .	5
2.2 Sentiment Analysis . . . . .	5
2.3 Machine learning . . . . .	7
2.3.1 Machine Learning Algorithms . . . . .	8
2.3.1.1 Supervised Learning . . . . .	8
2.3.1.2 Unsupervised Learning . . . . .	15
2.3.2 Word Representation . . . . .	18
2.3.3 Overfitting and Underfitting . . . . .	19
2.3.3.1 Overfitting . . . . .	19
2.3.3.2 Underfitting . . . . .	19
2.3.3.3 Bias vs. Variance . . . . .	20
2.3.4 Optimization algorithms . . . . .	21
2.3.4.1 Normal Equation . . . . .	21
2.3.4.2 Gradient descent . . . . .	21
2.3.4.3 Evaluation Metrics . . . . .	21
2.3.5 Deep Learning . . . . .	23
2.3.5.1 Deep Learning vs. Machine Learning . . . . .	23
2.3.5.2 Software and Frameworks . . . . .	26
2.3.5.3 Word Embedding . . . . .	27
2.3.5.4 Backpropagation . . . . .	29

<b>3</b>	<b>Literature Review</b>	<b>32</b>
3.1	Challenges for Arabic language sentiment analysis . . . . .	32
3.2	Arabic Dataset and Lexicon Generation . . . . .	33
3.3	Machine Learning approaches in Sentiment Analysis . . . . .	35
3.4	Deep Learning approaches in Sentiment Analysis . . . . .	37
<b>4</b>	<b>Implementation</b>	<b>48</b>
4.1	Used Datasets . . . . .	48
4.2	Initial Configuration . . . . .	51
4.2.1	Pre-processing Steps . . . . .	51
4.2.2	Hyper-parameter Configuration . . . . .	52
4.2.3	Word Embeddings Used . . . . .	52
4.2.4	Initial Architecture . . . . .	53
4.2.5	Language and Libraries Used . . . . .	53
4.3	Loading the data . . . . .	54
4.4	Architectures and Algorithms . . . . .	55
4.4.1	CNN with merging the datasets . . . . .	55
4.4.2	Hyper-parameter Tuning (Filter Size) . . . . .	57
4.4.3	CNN-SVM . . . . .	57
4.4.3.1	CNN-SVM with Merged datasets . . . . .	57
4.4.4	LSTM . . . . .	57
4.4.5	Multi-Channel CNN . . . . .	58
4.4.5.1	Multi-Channel CNN with merged datasets . . . . .	59
4.4.6	CNN-LSTM . . . . .	59
4.4.7	SVM with Merged Datasets . . . . .	59
4.4.8	Naïve Bayes with Merged Datasets . . . . .	60
4.5	Running Time . . . . .	65
4.6	Results . . . . .	69
4.6.1	Benchmarks . . . . .	69
4.6.2	Two Class Classification . . . . .	71
4.6.3	Three Class Classification . . . . .	71
4.7	Findings . . . . .	78
<b>5</b>	<b>Conclusion and Future Work</b>	<b>84</b>
5.1	Conclusion . . . . .	84
5.2	Future Work . . . . .	85
	<b>References</b>	<b>86</b>
	<b>Appendix A Appendix</b>	<b>95</b>



# List of Tables

2.1	Gradient Descent vs. Normal Equation . . . . .	21
2.2	CNN vs. RNN . . . . .	26
4.1	Used Datasets . . . . .	49
4.2	Polarity Percentage in Datasets . . . . .	50
4.3	First Merged Datasets Method Statistics . . . . .	56
4.4	Second Merged Datasets Method Statistics . . . . .	56
4.5	Experiments Running Time . . . . .	65
4.6	Two Class Classification Benchmarks . . . . .	69
4.7	Three Class Classification Benchmarks . . . . .	71
4.8	Two Class Classification Results . . . . .	72
4.9	Three Class Classification Results . . . . .	73
4.10	Best Algorithm for each dataset . . . . .	79
4.11	Experiment using New Merged Datasets . . . . .	81
4.12	Running Time of Merged Datasets Experiments . . . . .	82
4.13	Best Experiments using New Merged Datasets . . . . .	83

# List of Figures

2.1	Sentiment Analysis techniques . . . . .	6
2.2	Traditional Programming vs. Machine Learning . . . . .	7
2.3	Linear Regression . . . . .	9
2.4	Linear Classification . . . . .	10
2.5	Non-linear Classification . . . . .	10
2.6	Logistic Regression - Sigmoid Function . . . . .	11
2.7	SVM . . . . .	11
2.8	Decision Tree . . . . .	12
2.9	Random Forest . . . . .	13
2.10	Artificial Neural Network (ANN) . . . . .	13
2.11	K-Nearest Neighbors Example . . . . .	14
2.12	K-Means Algorithm . . . . .	15
2.13	Reduction of data from 2D to 1D . . . . .	16
2.14	Reduction of data from 3D to 2D . . . . .	16
2.15	Anomaly Detection . . . . .	17
2.16	Recommender Systems . . . . .	18
2.17	Overfitting vs. Underfitting . . . . .	20
2.18	Bias-Variance Tradeoff . . . . .	20
2.19	Confusion Matrix . . . . .	22
2.20	Performance of Deep Learning and Machine Learning . . . . .	23
2.21	Feature Engineering in Deep Learning and Machine Learning . . . . .	23
2.22	Example of a Deep Neural Network . . . . .	24
2.23	Example of RNN architecture . . . . .	24
2.24	Example of LSTM architecture . . . . .	25
2.25	Example of CNN architecture . . . . .	25
2.26	Word Embedding in space . . . . .	27
2.27	Examples for word embedding similarity detection . . . . .	28
2.28	CBOW model . . . . .	28
2.29	Skip-gram model . . . . .	29
2.30	Gradient Descent . . . . .	30
2.31	SGD without momentum . . . . .	30
2.32	SGD with momentum . . . . .	30
3.1	CNN Architecture used by Zhang et al. . . . .	40
3.2	CNN-SVM used by cao et al. . . . .	41
3.3	CNN Architecture used by Kim et al. . . . .	42
3.4	C-LSTM used by Zhou et al. . . . .	43
3.5	CNN Architecture used by Zhang et al. . . . .	43
3.6	MGNC-CNN used by Zhang et al. . . . .	44
3.7	CNN Architecture used by Svensson et al. . . . .	46
3.8	Convolutional Recurrent Neural Network (CRNN) Architecture used by Gonzalez et al. . . . .	47

4.1	Initial CNN Architecture . . . . .	53
4.2	LSTM Architecture . . . . .	57
4.3	LSTM Architecture after adding dropout layer . . . . .	58
4.4	Architecture using Filter sizes=[2,3,5] . . . . .	61
4.5	Architecture using Filter sizes=[2,3,5,7] . . . . .	62
4.6	Multi-Channel CNN Architecture . . . . .	63
4.7	CNN-LSTM Architecture . . . . .	64
4.8	The number of datasets that exceed the benchmark for each of our per- formed experiments for two-class . . . . .	79
4.9	The number of datasets that exceed the benchmark for each of our per- formed experiments for three-class . . . . .	80

# Abstract

Nowadays, people express their opinions, reviews on products, movies, hotels, restaurants...etc publicly on the internet on social media platforms, blogs or forums. The number of Arabic speaking users on the internet has been increased in the last decade and the research for analyzing the Arabic text has gained a lot of attention.

In my work, deep learning techniques are used to classify Arabic tweets and reviews into two classes (positive and negative) and three classes (positive, negative and neutral). Also, this work investigates whether deep learning can overcome ordinary machine learning algorithms and replace the effort of feature engineering in previous work.

Finally, deep learning proved to have better results than machine learning techniques for most of the used datasets by using a data augmentation architecture.

# Chapter 1: Introduction

In the last decade, the Internet has become one of the life essentials. People use the internet to do almost everything in life: paying bills, online shopping, hotel or restaurant reservation, jobs' application. . . etc. With the growth of the internet use, the development of smart phones and the availability and diversity of social media platforms has increased like Facebook, Twitter, Instagram. . . etc.

Nowadays, people become addicted to social networks where they can express their feelings, opinions, feedback and thoughts. For example, in Twitter, 6000 short messages (tweets) are posted by users every second.<sup>1</sup> [114]

As a result of the users' behavior towards social networks, researchers and companies began to think how to benefit from this freely available data by checking the feedback on their products, marketing some new items, classifying the audience based on opinions. At this point, sentiment analysis got a lot of attention in the field of data mining and natural language processing.

Research has focused on the field of natural language processing and how to analyze human language either verbal or written to extract useful information and to be used in other applications like sentiment analysis, speech recognition, machine translation and a lot more. Sentiment analysis is the task of analyzing and processing the text and detecting whether it expresses a positive or a negative opinion and in other cases detecting the polarity towards a specific topic. Sentiment analysis can be used by companies to gather users' reviews and opinions regarding certain product or service and take decision like continue to produce it or not.[117] [53]

Sentiment analysis is performed by the computers using machine learning and deep learning techniques as will be shown in the next section.

## 1.1 Problem Formulation

Machine learning is widely used for sentiment classification and recently, deep learning techniques showed a very good performance in various natural language processing (NLP) tasks. Deep learning is considered a part of machine learning methods. It is based on artificial neural networks consisting of an input layer and an output layer with many hidden layers between them where the name deep is derived. The role of these hidden layers is to extract features from input data automatically without manual feature engineering used in ordinary machine learning algorithms. Deep Learning has different techniques such as Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM).

Deep Learning has gained a huge success in the field of image recognition and it began to be widely used in natural language processing and specifically sentiment analysis. Many research papers have been published in the field of sentiment analysis either using machine learning techniques like SVM, Naïve bayes, KNN. . . etc or deep learning techniques like CNN and RNN. Most of these researches were analyzing the English language which is the most used language across the world.

---

<sup>1</sup><http://www.internetlivestats.com/twitter-statistics/>

The goal of this thesis is to implement different deep learning techniques with different architecture to evaluate the impact of variations of hyper-parameters and the network architecture on the overall performance. We used tweets and reviews from different websites for our datasets and we performed classification into positive and negative sentiments in addition to including the neutral class.

## 1.2 Motivation

Sentiment analysis is commonly used for English language while work on Arabic language is still in the early stages. Therefore, there is not too many Arabic resources (datasets and lexicons) when compared to the English language. However, analyzing the Arabic language has been the focus of many researchers in the past few years due to the increase of Arabic speaking users on the internet and specifically on social media forums.

Sentiment Analysis in Arabic is considered a challenging task because of two reasons. First, the Arabic language has a very complex morphology and structure. Second, the Arabic language has many variations such as the Classical Arabic (CA) which is mainly used in the Holy Quran, the Modern Standard Arabic (MSA) which is a modernized and simplified version of CA and Dialectal also known as colloquial Arabic (DA) where each region has its own dialect. MSA is understood by all Arabic countries due to its use in education, media and formal communication.

As stated in [46] and as per November 2015 by Internetworldstats<sup>2</sup>, the number of people speaking Arabic as first language exceed 267 million and the number of people speaking Arabic as a second language exceeds 250 million. For the Arabic Internet users, there are around 168.1 million users with a user growth rate of 6,592.5. These statistics makes research of sentiment analysis on the Arabic language important. [46]

## 1.3 Thesis Contribution

In this work, machine learning and deep learning techniques with various architectures are used to classify Arabic tweets and reviews into two classes (positive/ negative) and three classes (positive/ negative/ neutral).

Convolutional Neural Network and Long Short Term memory networks have been implemented with different input like a small dataset and augmented dataset also called in this thesis Merged datasets. We also studied the impact of different architectures like changing number of layers and using multi-channel CNN. In addition, we used two different pre-trained word embeddings to build our input for the deep learning networks.

For the machine learning algorithms, we used the SVM and Naïve Bayes algorithms as proved to be the best classification techniques in the field of sentiment analysis.

Finally, this work investigates whether deep learning can overcome ordinary machine learning algorithms and replace the effort of feature engineering and compare both techniques in terms of performance or accuracy obtained.

---

<sup>2</sup><http://www.internetworldstats.com/stats7.html>

## **1.4 Thesis Organization**

The rest of the paper is structured as follows. Chapter 2 gives a theoretical background on sentiment analysis techniques, machine learning algorithms for supervised and unsupervised learning and the metrics used for evaluation; this chapter also includes the explanation of neural networks and backpropagation and finally discussion on deep learning. Chapter 3 presents the literature review and the work done by previous research in the field of machine learning and deep learning specifically for sentiment analysis in both English and Arabic languages. Chapter 4 presents the main contribution in the thesis by explaining the experiments we performed and the results obtained. Finally, we discuss the conclusion and future work in Chapter 5.