



Faculty of Graduate Studies for Statistical Research
Department of Computer Sciences



A proposed Machine Learning Approach For Tracking The Doing of public employee Over Social Network

Prepared by
Suzan Hamed

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Master Degree of Computer Science

Supervised by

Prof. Hesham A. Hefny

Dr. Mostafa Ezaat



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition

Social Media (SM)

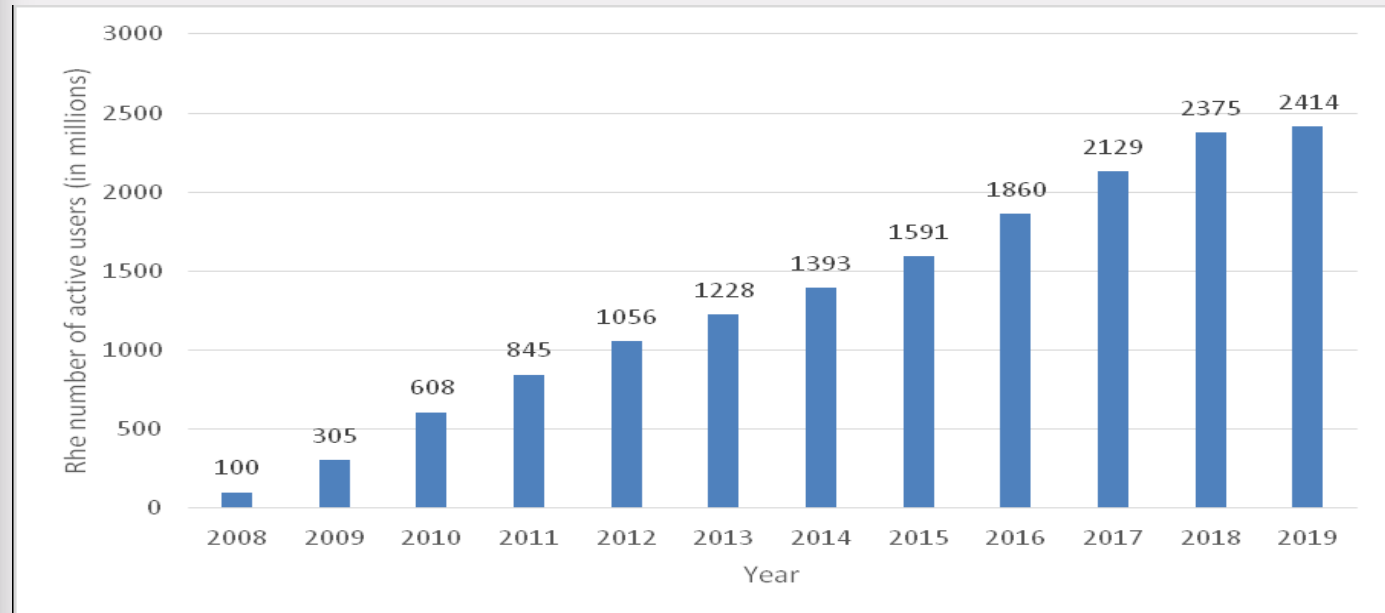
- ❑ **SM** users discuss everything about their lives, they care about Facebook (**FB**) posts, and tweets. They are sharing their opinions and showing their interest in real life.
- ❑ **SM** are considered an excellent source of information and can provide opinions, thoughts, and insights toward various important topics.
- ❑ **The Arabic content** on the internet has increased in volume, especially after the evolution of social networks. On **FB**, there are more than **7.5 Million** Arabic users who produce more than **11.8 million** comments per day.



introduction

Social Media (SM) (Cont.)

- ❑ **FB** is one of the most popular, widely used social media.
- ❑ **Official FB statistics** show that 240 million monthly users actively used Facebook in **2019**.



Number of monthly active Facebook users worldwide from 2008 to 2019.





Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition

Sentiment Analysis (SA)

- ❑ **SA** (also known as opinion mining, review mining, or Sentiment mining).
- ❑ **SA** is a field of Natural Language Processing (**NLP**).
- ❑ **SA** helps in achieving various goals like **observing public mood** regarding political topics, market intelligence, new product sales prediction, the measurement of customer satisfaction, and many more.



introduction

Sentiment Analysis (SA) (cont.)

- ❑ Hence, **SA** plays a significant role in our daily decision-making process.
- ❑ **SA** involves classifying opinions in text like "positive" "negative" or "neutral" or more (e.g., positive, negative, neutral, very positive, and very negative)  
- ❑ **SA** can be divided into **3 levels**, namely
 - ✓ document level,
 - ✓ sentence level, and
 - ✓ aspect level
(also known as word- or feature-level).





Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition

Deep learning (DL)

- ❑ DL is a subset of machine learning.
- ❑ Deep Neural Network (DNN) is essentially a neural network with a large number of hidden layers.
- ❑ DNNs attempt to simulate the behavior of the human brain.
- ❑ DNN Learn from large amounts of data.



introduction



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition



introduction

Deep learning(DL) (cont.)

- ❑ **DL** methods have the capability of feature representation by using supervised / unsupervised learning.
- ❑ **DL** is currently used in image applications, big data analyses, speech recognition, machine translation.
- ❑ **DL** need to use big data along with high computing performance to deal with the data.



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition

Types of Deep learning techniques

- ❑ **DL** techniques includes several networks architecture.
- ❑ **Two** of the most commonly used networks are:
 - Convolutional Neural Networks (**CNN**)
 - Recurrent Neural Network (**RNN**)
 - ✓ Long Short-Term Memory (**LSTM**) is a very effective type of **RNN**



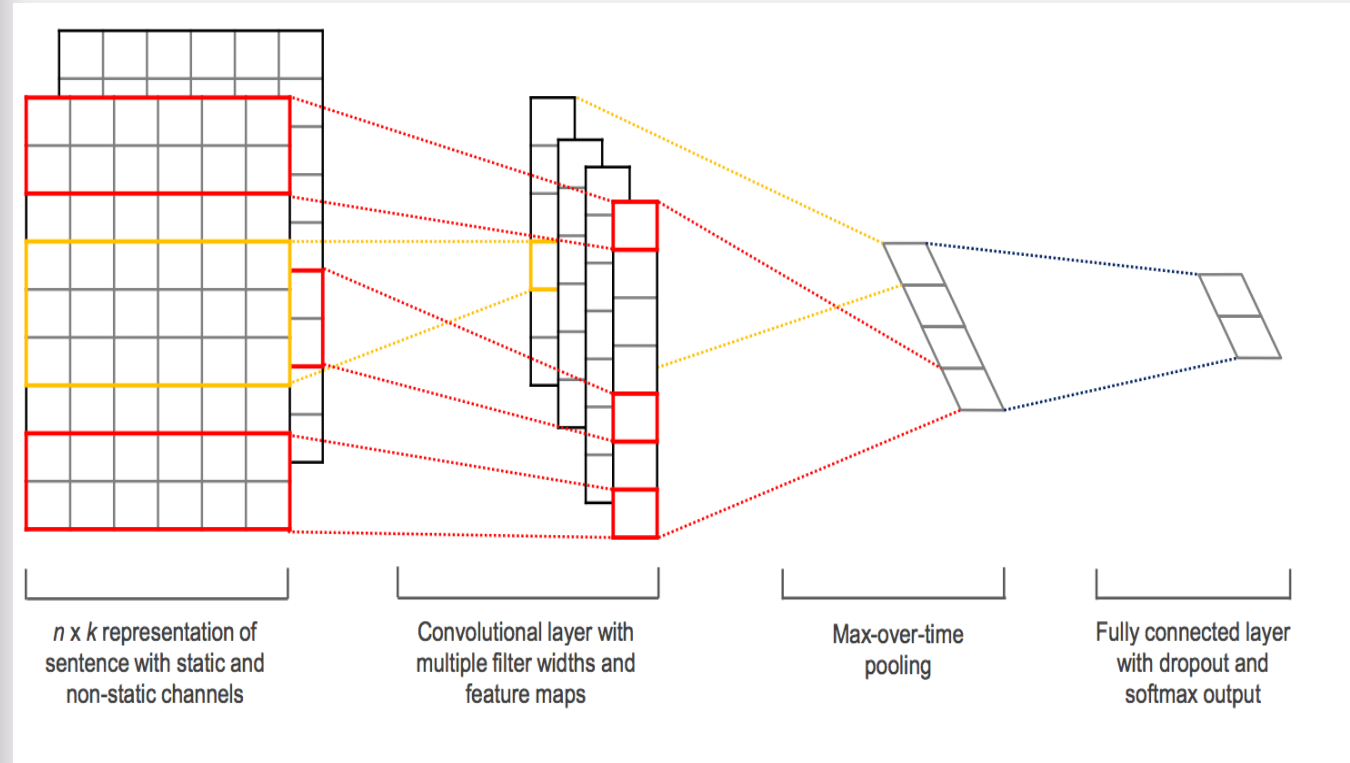
introduction

Convolutional Neural Network (CNN)

- ❑ CNN is a special kind of feed-forward neural network.
- ❑ CNN is inspired by the human visual cortex.
- ❑ CNN is used in speech recognition and computer vision.
- ❑ CNN is used in NLP for various text classification tasks.
- ❑ CNN has 3 main types of layers are convolutional, pooling, and fully-connected.
- ❑ CNN is able to handle data with high dimensions.
- ❑ CNN uses filters that move across the input layer and are used to calculate the values in the next layer.



Convolutional Neural Network (CNN)





Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem Definition

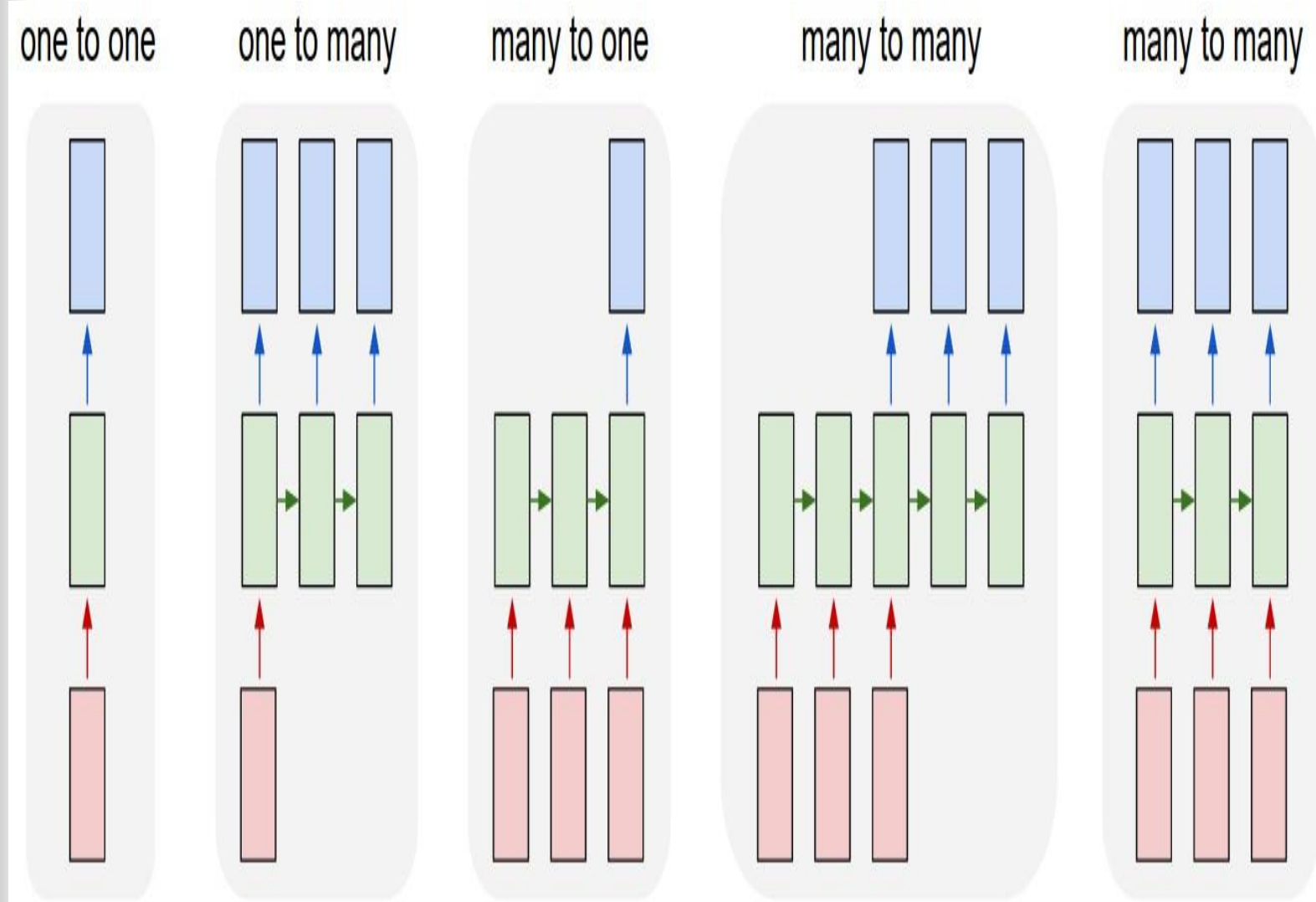
Recurrent Neural Networks (RNN)

- ❑ All the inputs in **RNN** are related to each other.
- ❑ **RNN** is suited for sequential inputs like speech and text and generating sequence.
- ❑ Main problem in **RNN** is that it is the gradient vanishing or exploding problems.
- ❑ **RNN** fails to solve problems that require learning long-term temporal dependencies (no memory)



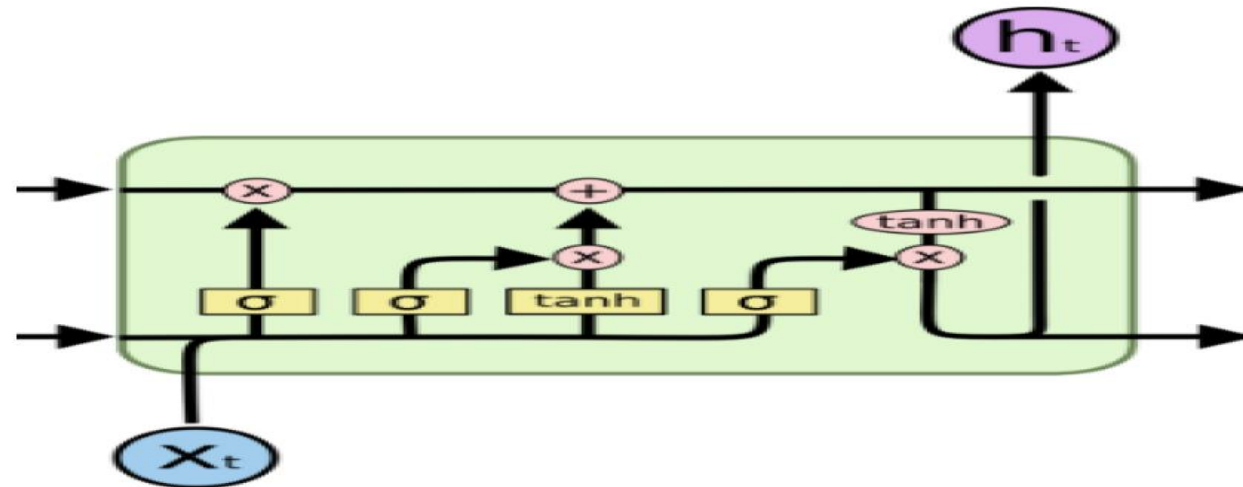
introduction

Types of (RNN)



Long Short -Term Memory (LSTM)

- ❑ **LSTM** is a special type of **RNN** capable of learning long-term sequences.
- ❑ **LSTM** architecture consists of input gates, output gates, forget gates, and cell memory.
- ❑ **LSTM** is more accurate on a larger dataset.



Text representation techniques

- ❑ **Word embedding** is the **key** concept used in **DL** for finding the vector value of words.
- ❑ **Word embedding** is a continuous space vector representation.
- ❑ The **word vector** is used as an input of the **DNN**.

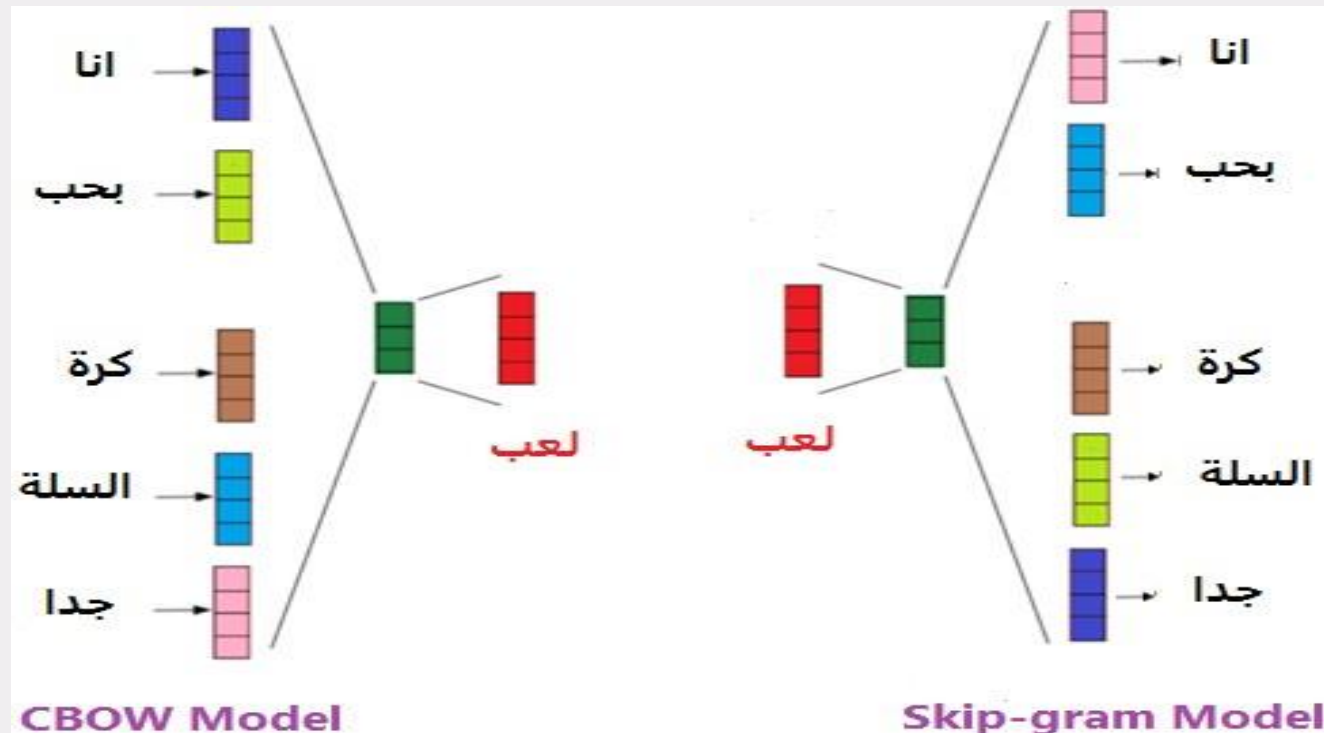
For example

- ✓ One hot encoding,
- ✓ N-grams,
- ✓ Word2vec
- ✓ Continuous Bag-of-Words (**CBOW**) model,
- ✓ Skip Gram model.



Text representation techniques (Cont.)

- ❑ Continuous Bag-of-Words model (CBOW) predicting the word given its context.
- ❑ Skip-gram model predicting the context given a word.



CBOW and Skip-gram models architecture



Sentiment Analysis for Arabic Language

Arabic Contents

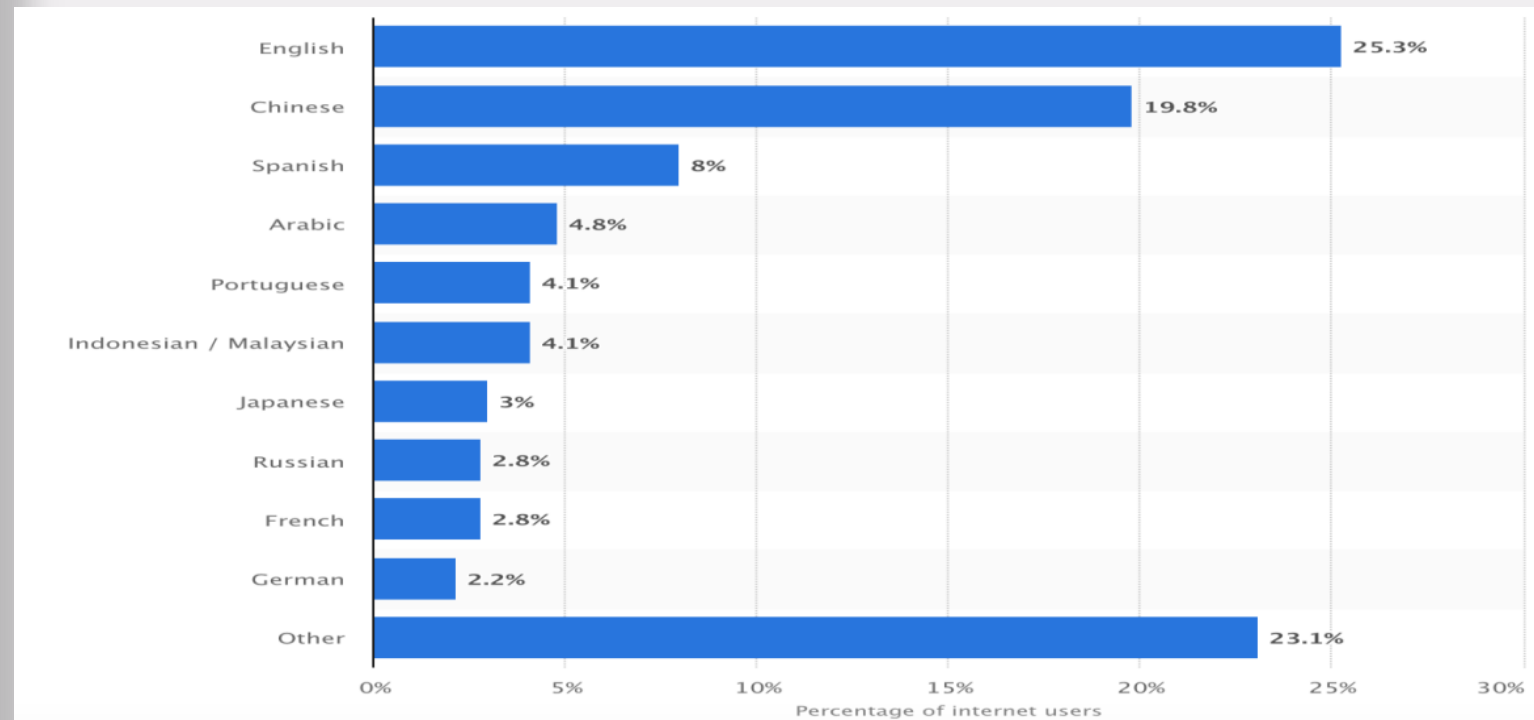
- ❑ Arabic language is one of six official languages of the United Nations.
- ❑ Arabic is the official language of 22 countries and is the major language in several areas of the world.
- ❑ Arabic language is classified into three types;
 - Classical Arabic ("AL-Qur'an")
 - Modern Standard Arabic (formal communications, television, radio, news, education)
 - Colloquial Arabic (shopping, chatting, or in their homes.)



Sentiment Analysis for Arabic Language

Arabic Contents (Cont.)

- Arabic is currently ranked as the 4th language used in the web, and there are about 178 million of Arabic Internet user .



Top 10 Languages used on the internet today



introduction

- 18

Problem Definition

- ❑ Although Arabic is one of the languages used in social networks, especially colloquial language, Arabic **SA** is quite limited due to the following reasons:
 - First, the Arabic language has a very complex morphological structure.
 - Second, most Arab users would resort to using colloquial Arabic instead of Modern Standard Arabic (**MSA**).
 - Third, the Arabic language is considered to have no common dialect due to the differences in regions from one region to another, and from one country to another.
 - Fourth: the lack of Arabic text.



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

Previous Works in Sentiment Analysis Using Deep Learning



previous work

Problem definition

introduction

Previous work

References	Model-Used	Data Set	Results
[Abdul-Mageed and Diab, 2012]	SVM and NB Algorithms	MSA collected corpora from three different resources: Wikipedia (1,508 sentences), web forum (1,019 sentences), and Penn Arabic Treebank (2,855 sentences).	SVM achieved 72% accuracy and NB achieved 65% accuracy
[Kim, 2014]	CNN model.	9 English datasets were obtained from movie reviews.	CNN achieved an accuracy of 89.6% using the MPQA dataset.



previous work

Problem definition

introduction

Previous work (cont.)

[Aldayel and Azmi, 2016]	SVM algorithm.	1,103 Saudi dialect tweets.	F-measure and accuracy of 84% and 84.01% respectively.
[Al-Ayyoub, et al., 2016]	NB, SVM, KNN, and decision tree.	Constructed a small-size corpus from book Reviews.	SVM achieved 45.7% accuracy. DT achieved 40.2% accuracy. NB achieved 32.2% accuracy. KNN achieved 38.6% accuracy.



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

Previous work (cont.)

[Wang, et al., 2016]	LSTM model	4,491 Customer reviews. (This dataset is expressed in 8 different languages.)	LSTM achieved an accuracy of 74.3%.
[Hassan, et al., 2016]	LSTM model	9,447 of Bengali tweets.	LSTM achieved an accuracy of 87%.
[Alomari, et al., 2017]	SVM and NB algorithms	1,800 Jordanian tweets.	The SVM algorithm achieved the highest performance (achieved 88% accuracy).



previous work

Problem definition

introduction

Previous work (cont.)

[Baly, et al., 2017]	LSTM model	Two corpora each of size 1,200 tweets on Egyptian and Emirati dialects.	The highest accuracy achieved was in the first corpus with an accuracy of 70%.
[Alayba and Iqbal, 2017]	CNN and (SVM, NB, and Logistic Regression).	2,026 Arabic tweets that cover three topics in the health field.	SVM with 91.37% accuracy outperformed CNN with 90% accuracy.
[Chen, et al., 2018]	LSTM model	3,200 English tweets have emojis	LSTM achieved an accuracy of 70.7%

Previous work (Cont.)

[Vizcarra, et al., 2018]	CNN model	Two different datasets of Spanish tweets: 1. InterTASS (2,907 tweets) 2. General TASS (7,219 tweets)	CNN achieved an accuracy of 61.82% and 73.22%.
[Heikal, et al., 2018]	CNN and LSTM models.	3,315 Arabic tweets	CNN achieved an accuracy of 64.30% while LSTM achieved accuracy of 64.75%.

Previous work (Cont.)

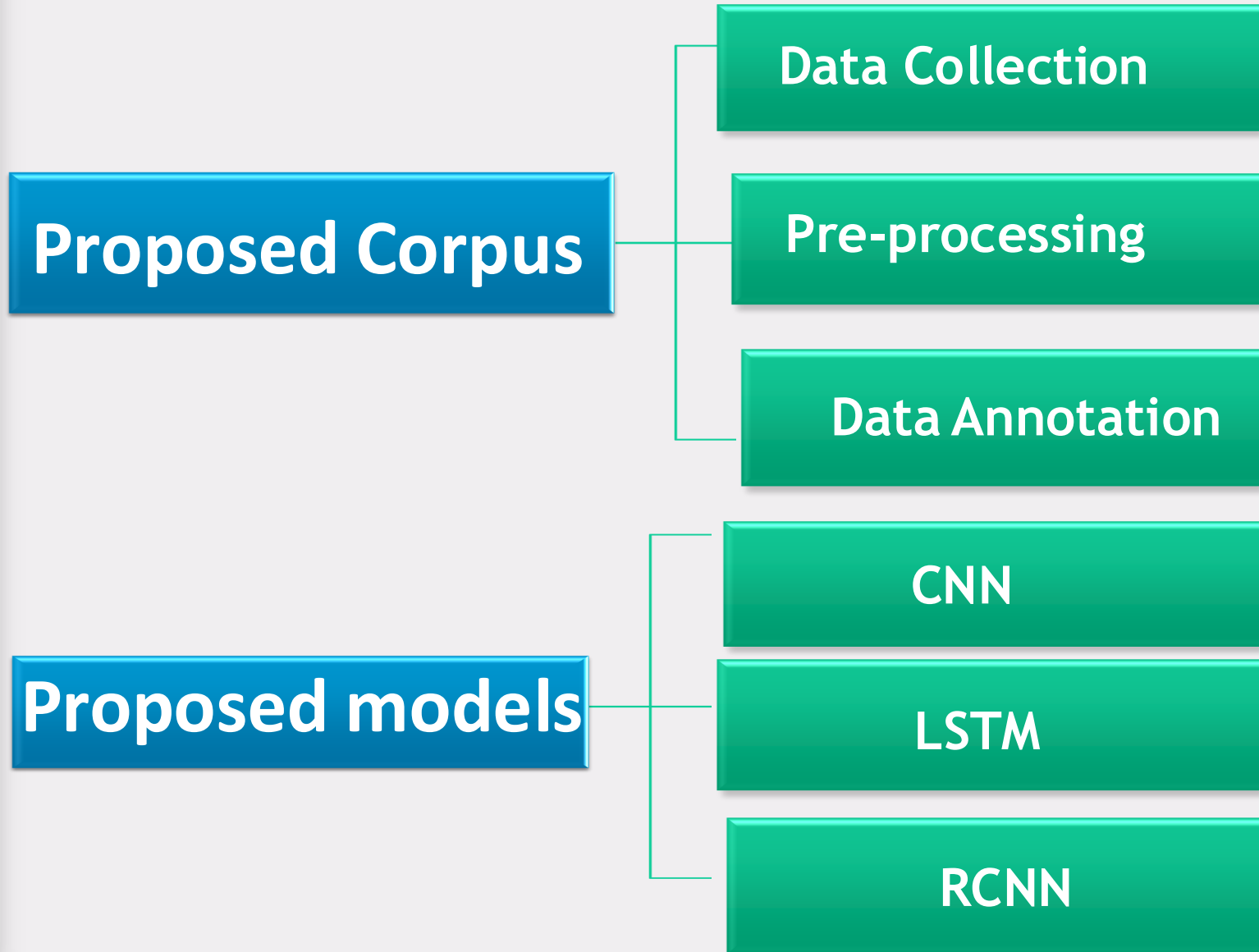
[Elzayady, et al., 2021]	CNN and LSTM models	Dataset HTL dataset.	CNN achieved an accuracy of 85.83%, while LSTM achieved an accuracy of 86.88%.
[Abdelhady, et al., 2022]	CNN and LSTM models	There are 10,000 tweets ArCovidVAC Dataset	LSTM achieved an accuracy of 79.29%, while CNN achieved an accuracy of 74.06%
[Alqarni & Rahman, 2023]	CNN and BiLSTM	There are three different corpora in three major regions in Saudi Arabia Arabic tweets with sizes of 25,860, 20,912 , and 9,613 which are: Riyadh, Dammam, and Jeddah ,respectively.	CNN achieved 92.80% accuracy while The performance of BiLSTM was scored at 91.99%

Research Objective

- ❑ The main objective of this research is to apply deep learning technology to sentiment analysis on social media in which slang Arabic dialect is used.
- ❑ Also, We are more concerned in getting as much Arabic material from which performance is evaluated through various methods of deep learning.



The Proposed Approach



Proposed Approaches

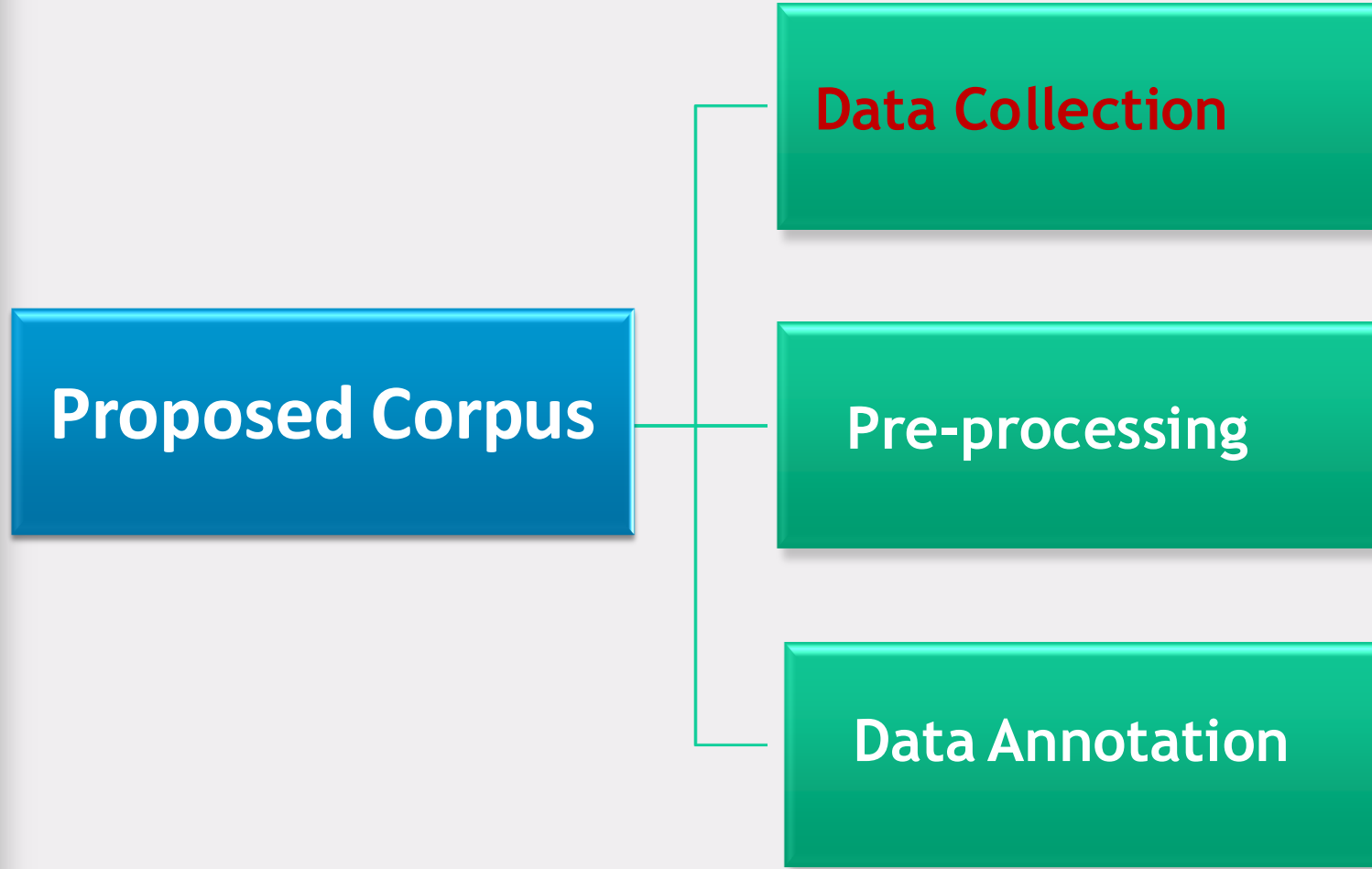
Objectives

previous work

Problem definition

introduction

The Proposed Corpus





Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Stage 1: Data Collection

Load Data (corpus)

□ comments were collected from APIs.

The collected comments span the period from November 24, 2019, to June 18, 2020. In this initial stage, a collection of 3,250,189 noisy raw comments was gathered.

Proposed Approaches

Objectives

previous work

Problem definition

introduction

C

Original comments

Proposed Corpus

Data Collection

Pre-processing

Annotation Data

Proposed Approaches

Objectives

previous work

Problem definition

introduction

The Proposed corpus (Cont.)



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Stage 2: preprocessing (Handling data)

- **Phase 1: Filtering Data**
- **Phase 2: Cleaning Data**
- **Phase 3: Normalization Data**

Proposed Approaches

Objectives

previous work

Problem definition

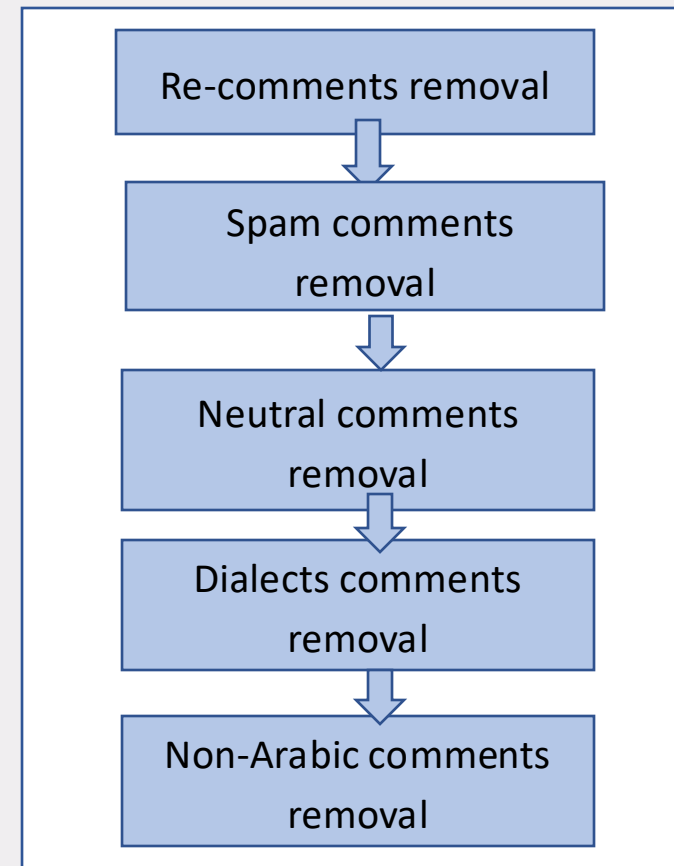
introduction

The Proposed corpus (Cont.)

Stage 2: preprocessing (Handling data) (Cont.)

Phase 1: Filtering Data

- ❑ After collecting the data from Facebook, we found that the collected comments have **many problems** that will affect the annotation process accuracy.
- ❑ This phase consisted of the shown **five** steps that are prepared **manually**.



comments filtering phase

The Proposed corpus (Cont.)

Stage 2: preprocessing (Handling data) (Cont.)

Phase 2: Cleaning Data

- ❑ This phase is important since it deals with the **noisy** nature of comments.
- ❑ data cleaning :
such as: (hashtag(#), username (@username), emoticons (=D, XD),
time stamps and URL).
- ❑ Some comments may contain:
(missing spelling words, missing letters, and written incorrectly).
- ❑ This **problem** is handled by **correcting** these **words** and wrong Characters.
- ❑ Some words are **combined** due to the space limitation of FB. This problem is handled by **adding spaces** between words.

The Proposed corpus (Cont.)

Stage 2: preprocessing (Handling data) (Cont.)

Phase 3: Normalization Data

- Normalization is the process of transforming the Arabic text in order to be **consistent**.

تا	تأ	تنا	تأ
----	----	-----	----

All these forms cause the word to be considered as three different words.

Normalization Rules

Rule	Example
Tashkeel	< - تُقَرَّب تقرب
Repeate	< يكررررر يكرر
Tatweel	> - هـلا الله
Hamza	> - (ء، عى، وء) ء
Alef	> ا (إ، أ، آ)
Lamalef	> (لا، لآ، لا) لا
Yeh	> ي (ى، ي)
Heh	> ه (ة، هـ)

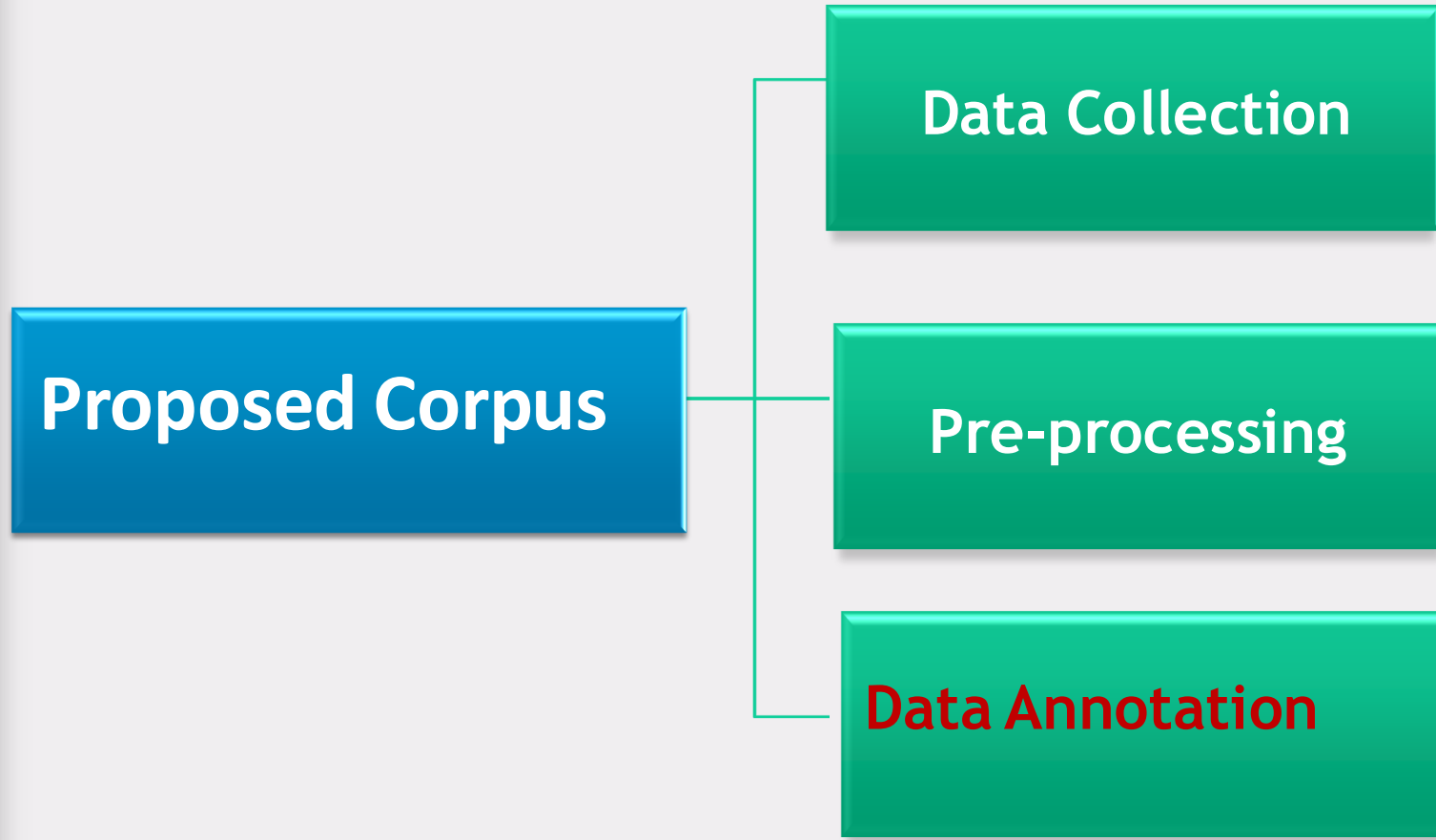


Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)



Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Stage 3: Data Annotation

- ❑ At this phase, the annotation process classifies each comment as “*positive*” or “*negative*”.

Proposed Approaches

Objectives

previous work

Problem definition

introduction

The Proposed corpus (Cont.)

Stage 3: Data Annotation (Cont.)

- ❑ New Arabic corpus for mixed comments written in **Egyptian dialect** and **MSA**.
- ❑ Cover **several topics** including:
 - ✓ Political, Social topics.
 - ✓ Users' opinions concerning different governorates.
- ❑ This corpus consists of **60000** comments classified as **30000** positive and **30000** negative.



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



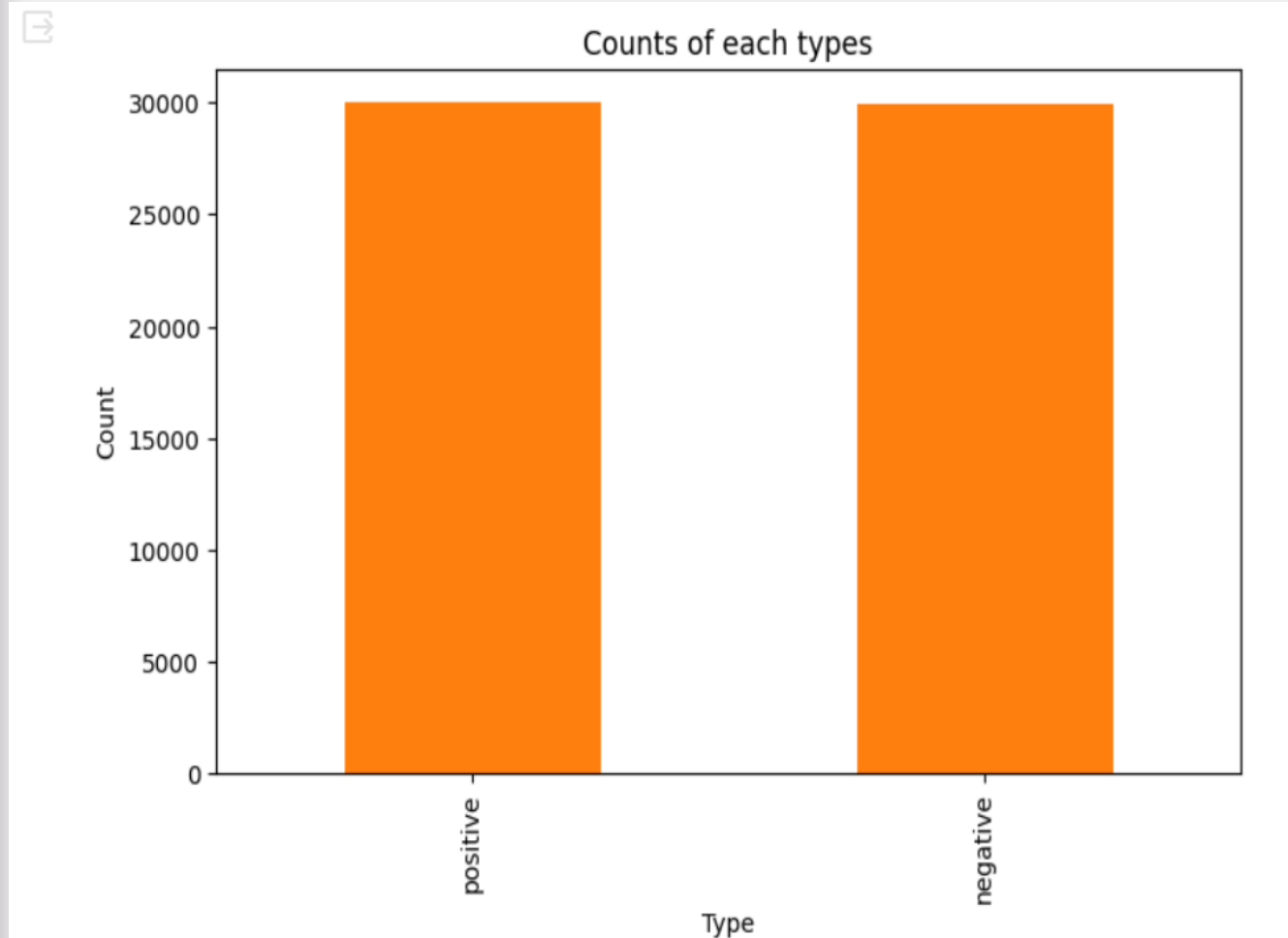
Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Stage 3: Data Annotation (Cont.)



Total distribution of Corpus (balanced data)

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Stage 3: Data Annotation (Cont.)

- It is divided into two columns:
 - The first column is called **comments**.
 - The second column is called **label**.

Proposed Approaches

Objectives

previous work

Problem definition

introduction

The Proposed corpus (Cont.)

	comments	label
38341	القهاوي شغاله في المنيب شارع المجزر الالي متساين ليه طبعاً عشان بيدفعوا لكم شهر يا حسبي اللو نعم الوكيل فيكم انتوا ترتشوا واحنا منعرفش ندخل ولا نخرج	negative
38342	معالي المحافظ المحترم شكرا علي كل اللي بتقدمه	positive
38343	الشوارع منكسره حاجه تقرف	negative
38344	لا حياه لمن تنادي حياه العيال خطر التكاك في كل حته حسبي الله	negative
38345	التكاك مبهدله الدنيا ولا حد بيعيرنا	negative
38346	منكم لله الشوارع منكسره	negative
38347	هااا انت المحافظ ده محترم قدموا بس شكوي	positive
38348	في حمله نضافه ورا شارع ترسا بجد شغل هائل	positive
38349	المياه بقالها قاطعه يومين حرام كدا	negative
38350	الماسوره اتصلحت عندنا شكرا لجهودكم	positive
38351	هاتوا لنا المياه ظلمة حسبي اللو نعم الوكيل حتي فالعيد قطعنها علينا ربنا المنتقم	negative
38352	عندنا المياه جت صلحوها ربنا يبارككم	positive
38353	ركبوا اضاءه تحت الكوبري من احسن حاجه اتعملت	positive
38354	المكان منور شكرا لمجهوداتكم	positive
38355	فعلا أصبحوا زى الوباء في منطقة مشعل وشوارعها واغراق الشوارع بالمياه من غسيل السيارات لتصبح برك مياه	negative
38356	ليه بتطولوا في تقديم الخدمات في حين بتخلصوها علي مزاجكم في يوم وليله	negative
38357	هانت شغالين فيها من الصبح ربنا يباركهم	positive
38358	ما فيش رقابه والمجاري طفحه	negative
38359	الشارع اتسلفت وبقي تمام في المربوطيه شكرا	positive
38360	الزباله في كل حته والشوارع بقت مليانه كلاب حرام عليكم	negative
38361	طعموا الكلاب وما قتلوهاش الدنيا لسه بخير	positive
38362	بالطريق بقي متسلفت حاجه عظمه	positive
38363	المصبيه اننا بنهاتي مع نفسنا ملفيش مسئول بيرد علينا حاجه مهزله	negative

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Corpus Statistics

Total number of comments	60000
Number of positive comments	30000
Number of negative comments	30000
Number of words	359,818
Max comment token	90
Number of tokens	1,953,869
Average tokens per comments	17

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed corpus (Cont.)

Word Embedding

- ❑ In very simplistic terms, Word Embedding is the text converted into vectors.
- ❑ The different types of word embeddings can be broadly classified into two categories:
 - Frequency-based Embedding
 - ✓ Count Vector
 - ✓ TF-IDF Vector
 - ✓ Co-occurrence Vector
 - Prediction-based Embedding
 - ✓ Word2vec
 - ✓ Continuous Bag-of-Words model (CBOW)
 - ✓ Skip-gram model

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

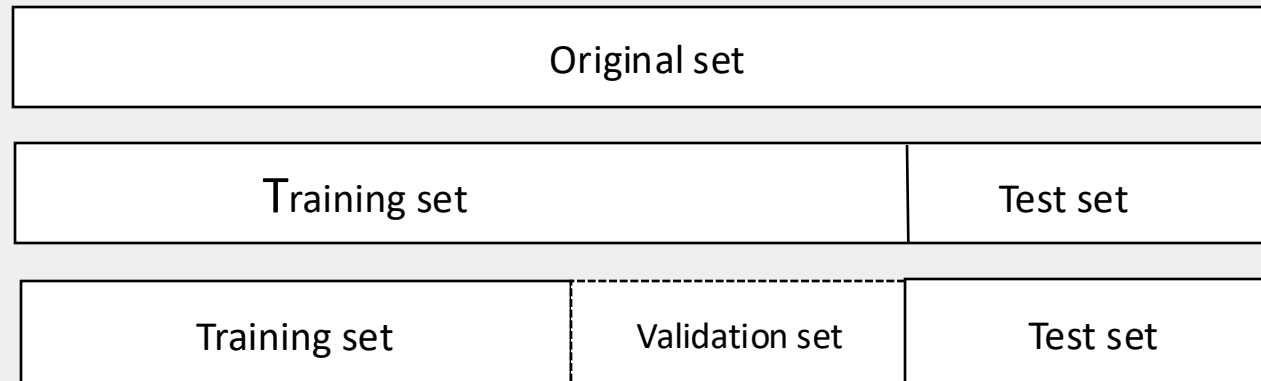
Results and Discussion

The Proposed corpus (Cont.)

Data Splitting

- ❑ It's split into training, validation, and test sets.
- ❑ create sets for the features and the labels.

train_x and train_y
test_x and test_y



Proposed Approaches

Objectives

previous work

Problem definition

introduction



The Proposed DNN Models

- ❑ Three commonly used **DNN** models are adopted in the proposed approach. They are:
 - Convolutional Neural Network (**CNN**)
 - Long short-term memory (**LSTM**).
 - Recurrent Convolutional Neural Network (**RCNN**)
- ❑ The first layer need to be prepared for each of the above three models is the **“Word Embedding Layer”**.



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed DNN Models (Cont.)

Embedding layer

Proposed Approaches

Objectives

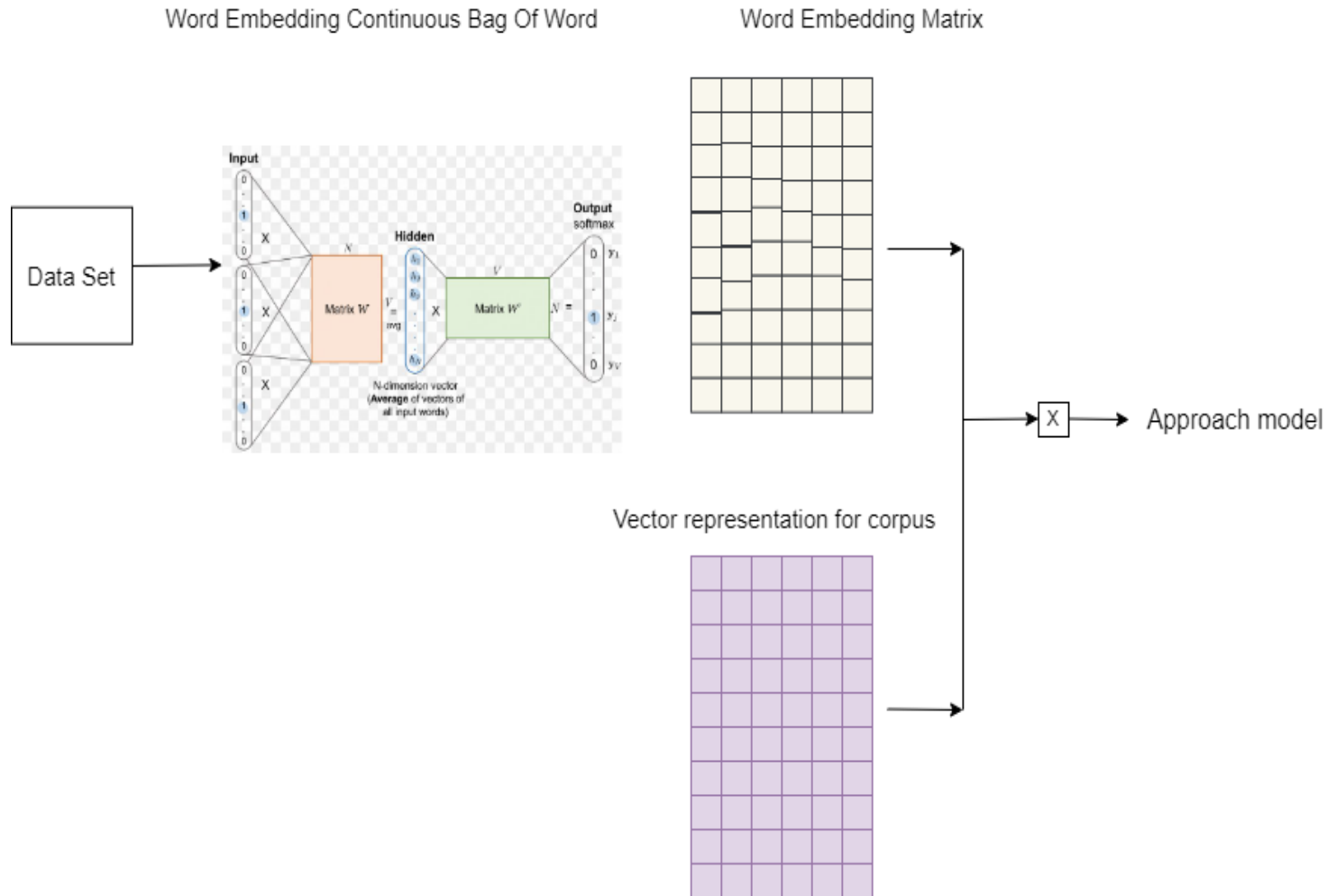
previous work

Problem definition

introduction

The Proposed DNN Models (Cont.)

Word Embedding using CBOW



Proposed Approaches

Objectives

previous work

Problem definition

introduction



Publications.

Conclusions & Future Work

Results and Discussion

The Proposed DNN Models (Cont.)

1. CNN Model

- ❑ Convolutional Neural Network (**CNN**)
- ❑ The **CNN** architecture consists of:
 - Embedding Layer,
 - Convolutional Layer,
 - Max-pooling Layer,
 - Flatten Layer,
 - Dense layer, and
 - Activation layer.

Proposed Approaches

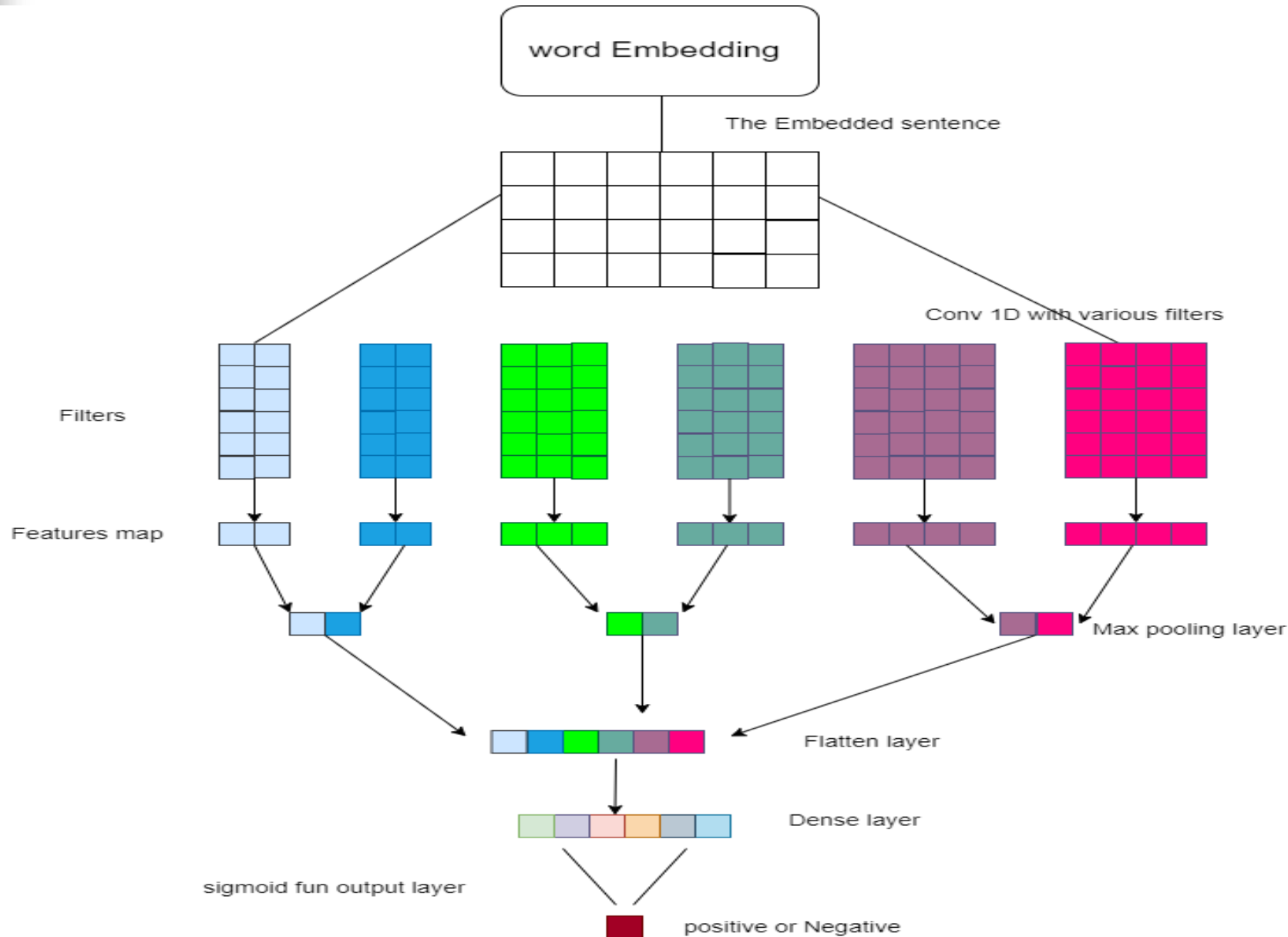
Objectives

previous work

Problem definition

introduction

Architecture of CNN Model



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



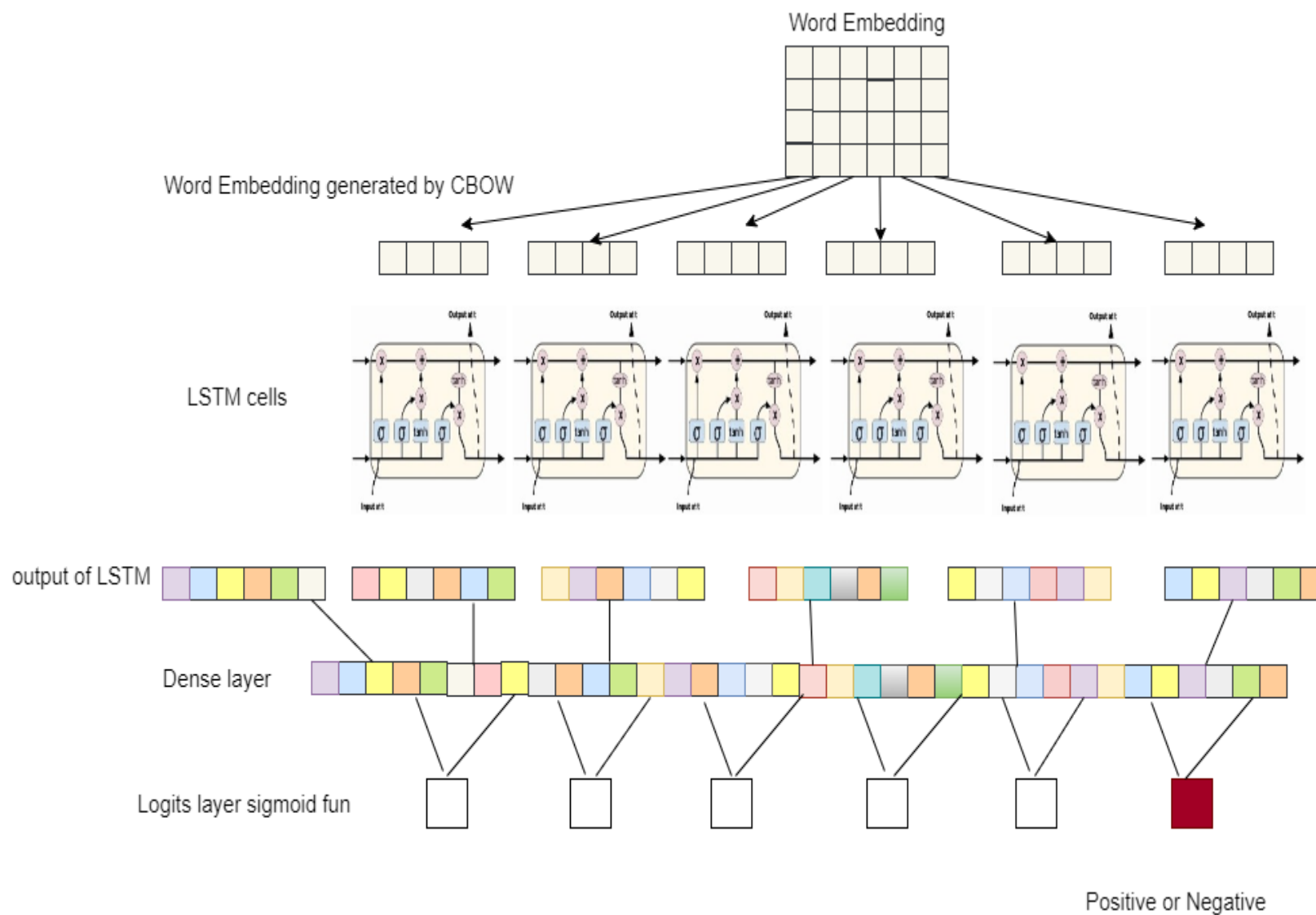
The Proposed DNN Models (Cont.)

2. LSTM Model

- ❑ Long short-term memory (**LSTM**).
- ❑ One of the advantages of the **LSTM** is the ability to capture the sequential data by considering the previous data.
- ❑ The **LSTM** architecture consists of:
 - Embedding Layer,
 - **LSTM** Layer,
 - Dense layer, and
 - Activation layer.



Architecture of LSTM Model

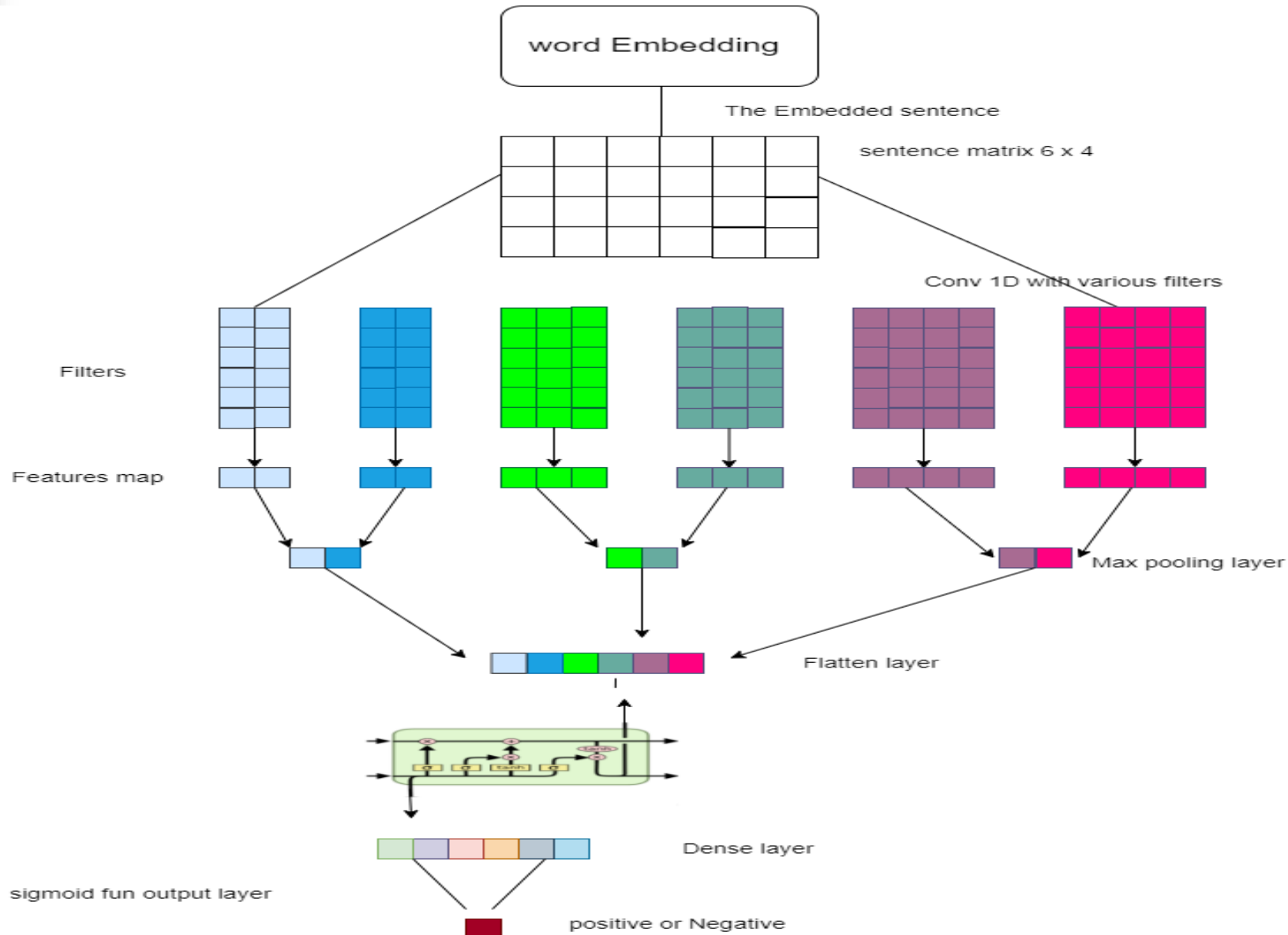




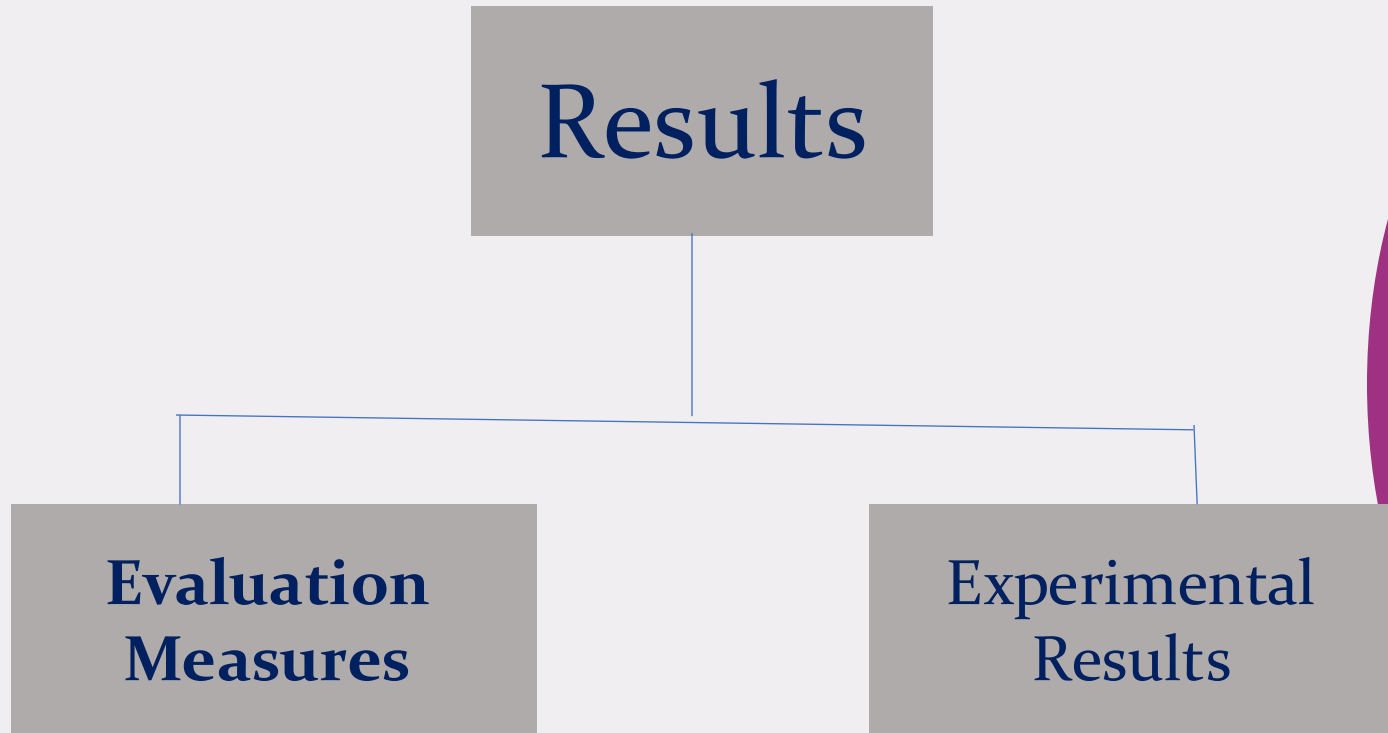
3. RCNN Model

- ❑ Recurrent Convolutional Neural Network (**RCNN**) is a network the **combining** the advantages of both **CNN** and **LSTM** neural networks.
- ❑ The **RCNN** architecture consists of:
 - Embedding Layer,
 - Convolutional Layer,
 - Max-pooling Layer,
 - Flatten Layer,
 - **LSTM** layer, and
 - Dense layer, and
 - Activation layer.

Architecture of RCNN Model



Evaluating The proposed Approach



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

Evaluating The proposed Approach (Cont.)

Evaluation Measures

	Actual positive (1)	Actual negative (0)
predicted positive (1)	true positive (TP)	false positive (FP)
predicted negative (0)	False negative(FN)	true negative(TN)

Accuracy = $(TP+TN) / (TP+TN+FP+FN)$

Precision= $TP/ (TP+FP)$

Recall= $TP/ (TP+FN)$

F1 score= $(2* Precision*Recall)/ (Precision+ Recall)$



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Evaluating The proposed Approach (Cont.)

- ❑ The experiment findings on the three suggested models are presented in this section.
- ❑ With ratios of **80%** to **20%**, **70%** to **30%**, and **60%** to **40%**, respectively, the dataset was split into training and testing subsets, with the remaining **10%** being utilized for validation.
- ❑ Before dividing the datasets for training and testing in **five** cross-validations, random shuffling was carried out.



Publications.

Conclusions & Future Work

Evaluating The proposed Approach (Cont.)

The experiment of the CNN Model

- ❑ It shows the average performance metrics for various data divisions using the **CNN** model.
- ❑ It also presents the typical performance metrics for different data divisions using the **CNN** model.

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Evaluating The proposed Approach (Cont.)

The experimental Results of the CNN Model (Cont.)

Average performance measures for various data divisions using the CNN model.

Data set	Performance Measures	Cross-validation #1	Cross-validation #2	Cross-validation #3	Cross-validation #4	Cross-validation #5
1	Average accuracy	78.74%	77.28%	78.65%	50.77%	79.27%
	Average recall	72.33%	75.61%	82.4%	53.65%	75.24%
	Average precision	83.78%	80.49%	78.71%	49.48%	83.09%
	Average f-score	76.58%	75.86%	79.74%	50.7%	77.69%
2	Average accuracy	78.71%	77.99%	77.87%	79.20%	78.25%
	Average recall	73.78%	78.48%	78.40%	77.75%	77%
	Average precision	82.85%	79.14%	79.62%	81.45%	80.8%
	Average f-score	77.40%	77.91%	77.69%	78.90%	77.65%
3	Average accuracy	78.38%	77.99%	78.69%	78.38%	77.68%
	Average recall	72.45%	78.48%	83.02%	76.02%	76.43%
	Average precision	84.27%	79.14%	77.31%	81.90%	80.47%
	Average f-score	76.64%	77.91%	79.57%	77.62%	76.63%

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

Introduction

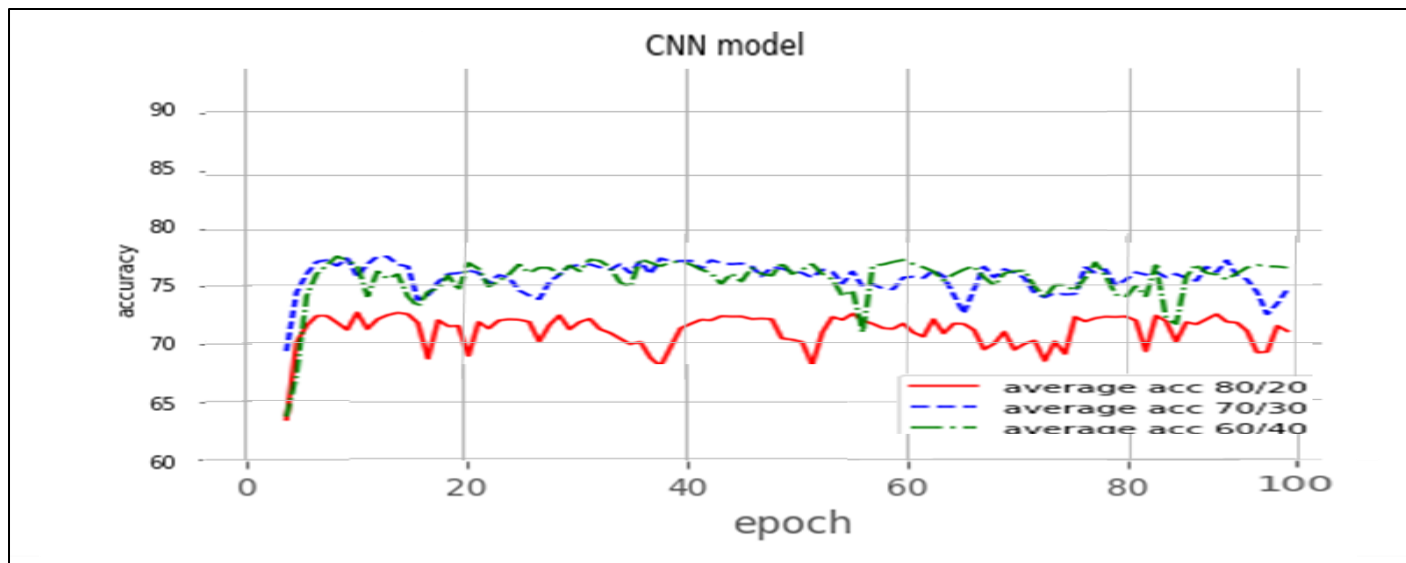


Evaluating The proposed Approach (Cont.)

The experimental Results of the CNN Model (Cont.)

Average performance measures for various data divisions using the CNN model.

Model	Data split		AVG Accuracy 5-fold	AVG recall 5-fold	AVG precision 5-fold	AVG f-score 5-fold
	Trainset	Testset				
CNN	80%	20%	72.94%	71.85%	75.11%	72.11%
	70%	30%	78.40%	77.08%	80.77%	77.8%
	60%	40%	78.22%	77.28%	80.62%	77.67%
Total AVG			76.52%	75.40%	78.83%	75.86%



Average accuracy per 100 epochs for various data divisions with CNN model

Evaluating The proposed Approach (Cont.)

The experiment of the LSTM Model

- ❑ It Displays the average performance metrics for different data divisions using the **LSTM** model.
- ❑ It also and the average performance metrics for different data divisions using the **LSTM** model.

Evaluating The proposed Approach (Cont.)

The experimental Results of the LSTM Model

Average performance measures for different data divisions using the LSTM model.

Data set	Performance Measures	Cross-validation #1	Cross-validation #2	Cross-validation #3	Cross-validation #4	Cross-validation #5
1	Average accuracy	81.92%	82.96%	82.21%	82.17%	82.2%
	Average recall	81.31%	82.75%	82.68%	82.09%	83.17%
	Average precision	82.67%	83.63%	82.58%	82.45%	81.67%
2	Average f-score	81.77%	83.01%	82.41%	82.06%	82.2%
	Average accuracy	82.10%	83.05%	82.36%	82.11%	82.05%
	Average recall	80.90%	82.25%	81.93%	81.29%	82.38%
3	Average precision	83.27%	83.76%	83.56%	83.09%	81.90%
	Average f-score	81.88%	82.82%	82.56%	81.97%	81.93%
	Average accuracy	82.08%	81.68%	81.23%	81.29%	82.30%
	Average recall	81.01%	81.8%	80.92%	80.27%	81.07%
	Average precision	83.28%	81.86%	82.57%	82.29%	83.31%
	Average f-score	81.96%	81.65%	81.55%	81.07%	82%

Publications.

Conclusions & Future

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

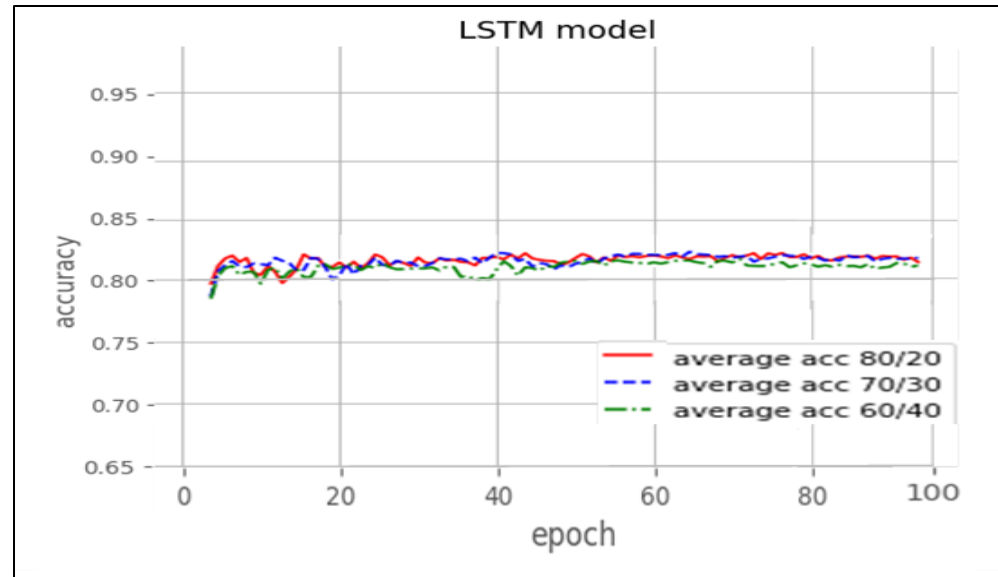


Evaluating The proposed Approach (Cont.)

The experiment of the LSTM Model (Cont.)

Average performance measures for different data divisions using the LSTM model.

Model	Data split		AVG Accuracy 5-fold	AVG (recall) 5-fold	AVG (precision) 5-fold	AVG f-score 5-fold
	Train set	Testset				
LSTM	80%	20%	82.29%	82.4%	82.6%	82.29%
	70%	30%	82.33%	81.75%	83.11%	82.23%
	60%	40%	81.71%	81.01%	82.66%	81.64%
Total AVG			82.11 %	81.72 %	82.79 %	82.05 %



Average accuracy per 100 epochs for various data divisions with LSTM model

Evaluating The proposed Approach (Cont.)

The experiment of the RCNN Model

- ❑ It displays the average performance measures for the **RCNN** model. Additionally.
- ❑ It also shows the typical performance measures for different data divisions using the **RCNN** model.

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

Evaluating The proposed Approach (Cont.)

The experiment of the RCNN Model (Cont.)

Average performance measures for different data divisions using the RCNN model.

Data set	Performance Measures	Cross-validation #1	Cross-validation #2	Cross-validation #3	Cross-validation #4	Cross-validation #5
1	Average accuracy	79.37%	79.71%	79.28%	78.97%	80.8%
	Average recall	78.11%	77.82%	77.86%	74.52%	78.63%
	Average precision	81.51%	82.27%	81.35%	83.01%	81.37%
	Average f-score	78.8%	79.35%	78.62%	77.25%	79.36%
2	Average accuracy	78.85%	79.06%	79.71%	79.16%	78.46%
	Average recall	78.92%	76.33%	82.26%	75.81%	79.67%
	Average precision	79.74%	82.08%	79.47%	82.65%	79.07%
	Average f-score	78.48%	77.97%	80.23%	78.10%	78.31%
3	Average accuracy	79.92%	79.79%	78.89%	77.97%	79.09%
	Average recall	79.74%	78.93%	76.31%	80.24%	77.71%
	Average precision	80.89%	81.26%	81.62%	78.43%	81.45%
	Average f-score	79.62%	79.44%	78.11%	78.08%	78.57%

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

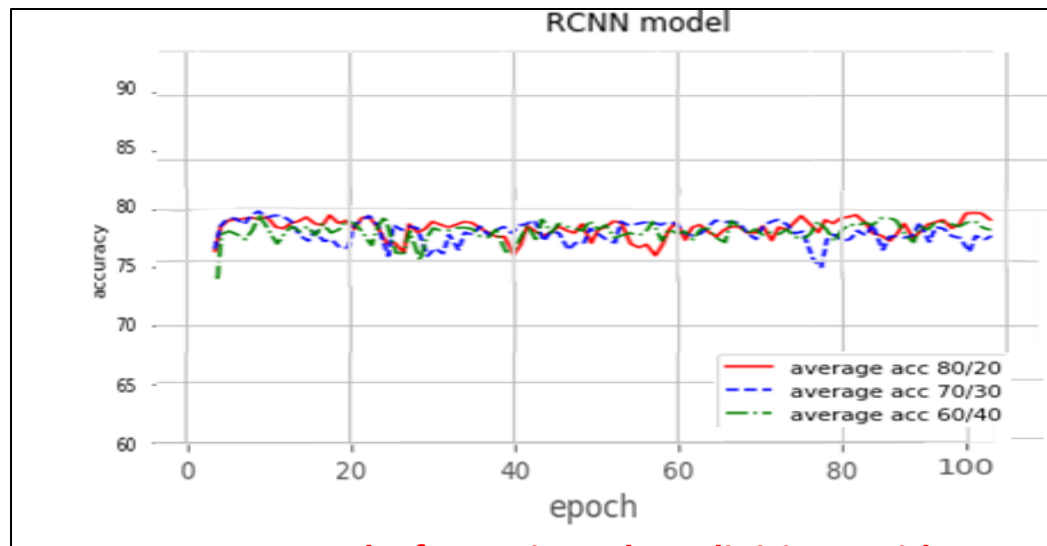


Evaluating The proposed Approach (Cont.)

The experiment of the RCNN Model (Cont.)

Average performance measures for different data divisions using the RCNN model.

Model	Data split		AVG Accuracy5- fold	AVG recall 5-fold	AVG precision 5-fold	AVG f-score5- fold
	Trainset	Testset				
RCNN	80%	20%	79.62%	77.38%	81.90%	78.61%
	70%	30%	79.04%	78.59%	80.60%	78.62%
	60%	40%	79.13%	78.58%	80.73%	78.76%
Total AVG			79.26%	78.18%	81.07%	78.66%



Average accuracy per 100 epochs for various data divisions with RCNN model

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Evaluating The proposed Approach (Cont.)

The experiment results after Data Augmentation Techniques

- ❑ This section presents the model's output after using the shuffling data augmentation method, which increases the number of comments in our corpus to 60,000 .
- ❑ The results show the average performance metrics for different data divisions using the **LSTM** model with data augmentation.
- ❑ Additionally, the results displays the average performance metrics for different splits using the **LSTM** model with data augmentation.

Evaluating The proposed Approach (Cont.)

The experiment results after Data Augmentation Techniques (Cont.)

Average performance measures for different data divisions using the LSTM model with data augmentation.

Data set	Performance Measures	Cross-validation #1	Cross-validation #2	Cross-validation #3	Cross-validation #4	Cross-validation #5
1	Average accuracy	89.54%	89.46%	89.32%	89.49%	89.78%
	Average recall	89.53%	88.97%	89.03%	89.55%	89.89%
	Average precision	89.14%	89.24%	89.22%	89.15%	89.43%
	Average f-score	89.29%	89.05%	89.08%	89.31%	88.97%
2	Average accuracy	88.67%	88.04%	89.48%	88.62%	88.59%
	Average recall	90.35%	89.72%	89.84%	90.17%	90.19%
	Average precision	86.92%	86.23%	87.08%	87.02%	86.92%
	Average f-score	88.54%	87.86%	88.36%	88.48%	88.46%
3	Average accuracy	88.8%	88.69%	88.47%	88.78%	88.06%
	Average recall	89.75%	89.40%	89.53%	90.07%	89.58%
	Average precision	87.41%	87.59%	87.24%	87.29%	86.40%
	Average f-score	88.5%	88.44%	88.33%	88.62%	87.93%

Publications.

Conclusions & Future

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

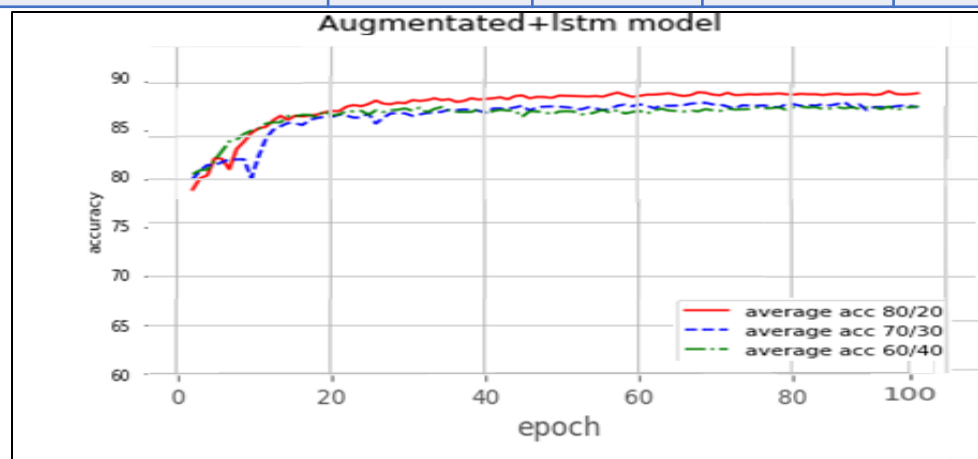


Evaluating The proposed Approach (Cont.)

The experiment results after Data Augmentation Techniques (Cont.)

Average performance measures were obtained for various splits using an LSTM model with data augmentation.

Model	Data split		AVG Accuracy 5-fold	AVG recall 5-fold	AVG precision 5-fold	AVG f-score5-fold
	Train set	Testset				
LSTM +AUG	80%	20%	89.51%	89.39%	89.23%	89.23%
	70%	30%	88.48%	90.05%	86.83%	86.54%
	60%	40%	88.56%	89.66%	87.18%	88.36%
Total AVG			88.85%	89.7%	87.74%	88.04%



Average accuracy per 100 epochs for various data divisions with augmentation data in LSTM model



Evaluating The proposed Approach (Cont.)

Comparison and Discussion

- ❑ In this section, the average performance measures of the **CNN**, **LSTM**, and **RCNN** models are compared.
- ❑ Also, it shows a summary and comparison of all the average performance metrics obtained after applying Data Augmentation methods to the **CNN**, **LSTM**, and **RCNN** models.

Evaluating The proposed Approach (Cont.)

Comparison and Discussion (Cont.)

A comparison between the average performance measures of the CNN, LSTM, and RCNN models is presented.

Models	AVG Accuracy (%)	AVG Recall (%)	AVG Precision (%)	AVG f-score (%)
CNN	76.52	75.40	78.83	75.86
LSTM	82.11	81.72	82.79	82.05
RCNN	79.26	78.18	81.07	78.66

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Evaluating The proposed Approach (Cont.)

Comparison and Discussion (Cont.)

A comparison is made between the average performance measures of the CNN, LSTM, and RCNN models after applying Data Augmentation.

Models	AVG Accuracy (%)	AVG Recall (%)	AVG Precision (%)	AVG f-score (%)
CNN	76.52	75.40	78.83	75.86
LSTM	88.85	89.7	87.74	88.04
RCNN	79.26	78.18	81.07	78.66



Publications.

Conclusions & Future Work

The performance of CNN and LSTM models of other previous works with considerably small data sets

Researcher Name	Model Used	Data Set	Results	Results the dependent model
[Baly, et.al 2017]	LSTM model	1,200 tweets on Egyptian.	achieved with an accuracy of 70%.	achieved with an accuracy of 73.8%.
[Heikal, et al. 2018]	CNN and LSTM models.	3,315 Arabic tweets	CNN achieved accuracy of 64.30% While , LSTM achieved accuracy of 64.75%.	CNN achieved accuracy of 65.7% While , LSTM achieved accuracy of 71.83%
[Abdelhady, et al., 2022]	CNN and LSTM models	There are 10000 tweets ArCovidVA C Dataset	LSTM achieved an accuracy of 79.29%, while CNN achieved an accuracy of 74.06%	LSTM achieved an accuracy of 82.78%, while CNN achieved an accuracy of

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction



Conclusion

- ❑ This thesis introduced a labeled corpus of 60k Arabic comments that discuss several domains of topics including political, sportive, health, people's opinions, and assessments about Giza and Cairo, written in colloquial Arabic.
- ❑ Additionally, it applied three deep-learning techniques to the proposed corpus.
- ❑ It also experimented with the performance of the corpus on **CNN**, **LSTM**, and **RCNN**.
- ❑ With the help of the word-embedding technique as an input layer to the three models, the **LSTM** achieved a high accuracy of 82.1% which outperformed **CNN** with an accuracy of **76.52%** and **RCNN** with an accuracy of **79.26%**.
- ❑ Also, by applying a data augmentation technique on the corpus, **LSTM** has improved further with an accuracy of **88.85%**.

Thesis Contribution

The main Contributions of this thesis are as follows:

- ❑ Building an annotated corpus of **60,000** comments covering people's opinions and assessments about Giza and Cairo, written in colloquial Arabic.
- ❑ The constructed corpus together with word embedding technique (**CBOW**) from the Cairo and Giza group have been adopted to train three DNNs models namely **RCNN**, **LSTM**, and **CNN**.
- ❑ Also, data augmentation technique has been adopted to improve the performance of the proposed approach.

Future Work

The future works related to this research can be summarized as follows:

- Applying other deep learning models on the corpus proposed in this study and comparing the results. In addition, **CNN**, **LSTM**, and **RCNN** models with the pre-trained skip-gram model on the proposed corpus and compared the results.
- Also, applying different types of data augmentation techniques to the deep learning models and the results.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

Publications

1. *Suzan Hamed, Mostafa Ezzat and Hesham Hefny,” A Review of Sentiment Analysis Techniques”, International Journal of Computer Applications 176(37):20-24, July 2020.*

The screenshot displays the homepage of the International Journal of Computer Applications (IJCA). The header features the journal's logo and name. A sidebar on the left lists navigation options such as Home, Archives, Special Issues, and Proceedings. The main content area highlights 'Most Read Research Articles' and a 'Call for Paper - August 2023 Edition'. The featured article, 'A Review of Sentiment Analysis Techniques', is shown with its title, authors (Suzan Hamed, Mostafa Ezzat, Hesham Hefny), and a DOI link (10.5120/ijca2020920480).

Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

Publications

2. Suzan Hamed, Mostafa Ezzat, Hesham Hefny, “Augmented Deep Learning Model for Social Network Sentiment Analysis”, *International Journal of Computer Applications Intelligent Systems and Applications in Engineering IJISAE* 10(4):246-255, 2022.

The screenshot shows the homepage of the International Journal of Intelligent Systems and Applications in Engineering (IJISAE). The header features the journal's logo, name, ISSN (2147-6799), and navigation links (Register, Login). The main content area displays the article title "Augmented Deep Learning Model for Social Network Sentiment Analysis" by Suzan Hamed, Mostafa Ezzat, and Hesham Hefny. A diagram of the proposed model is shown, illustrating the flow from word embeddings to sentiment prediction. The diagram includes a "word embedding corpus" at the bottom, which feeds into "word embedding generated by CBOW" and "word embedding generated by LSTM" blocks. These blocks feed into a "LSTM cell" block, which then feeds into a "Logits layer" block. The "Logits layer" block outputs "Positive or negative prediction".

International Journal of Intelligent Systems and Applications in Engineering
ISSN : 2147-6799

HOME / CURRENT / ARCHIVES / ANNOUNCEMENTS / ABOUT ▾

HOME / ARCHIVES / VOL. 10 NO. 4 (2022) / Research Article

Augmented Deep Learning Model for Social Network Sentiment Analysis

Suzan Hamed
Faculty of Graduate Studies for Statistical Research, Cairo University
<https://orcid.org/0000-0003-2819-6036>

Mostafa Ezzat
Associate professor, computer science department, faculty of Graduate Studies for Statistical Research, Cairo University
<https://orcid.org/0000-0001-5684-1878>

Hesham Hefny
Head of computer science department, faculty of Graduate Studies for Statistical Research, Cairo University
<https://orcid.org/0000-0001-5862-675X>

ANNOUNCEMENTS

Information for Authors

March 6, 2023
Information for Authors:
We are pleased to inform that we are now collaborating with **Elsevier Digital Commons** for much better visibility of journal. Further authors will be able to observe their citations, metric like PlumX from journal website itself. **IJISAE** will be in transition from **OJS** to **Digital Commons framework** in next few months so if there is any queries or delays contact directly on editor@ijisae.org

Activate V



Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction

Thank you



Questions ?

Publications.

Conclusions & Future Work

Results and Discussion

Proposed Approaches

Objectives

previous work

Problem definition

introduction