$[]shapes.geometric, arrows.meta, arrows, calc, intersections, matrix, positioning, patterns, decorations.text, matr$

xy

ABBREVIATIONS

ML
DL
NN
DNN
RNN
LSTM
BILSTM
GRU
BinE
OneE
TwoE
APCD
F1 Score
ReLU
CNN

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF FIGURES

$f(x) = x^2$ $\quad f(x) = x^2 \alpha_1 \alpha_2$

$\sigma(z) o tanh(z) Relu(x)$

$\{Trimming, NoTrimming\} \times \{Diacritics, NoDiacritics\} \times \{OneE, BinE, TwoE\}\{7L, 4L\} \times \{82U, 50U\} \times \{0W, 1W\} \times \{$
$\{Trimming\} \times \{Diacritics\} x Rigz$

LIST OF TABLES

ABSTRACT

binary one-hot two-hot binary one-hot

Chapter 1

INTRODUCTION

1.1 Thesis Outline

1.2 Arabic Poetry

$Al\text{-}Farahidi\; metersseas\; Al\text{-}Arud$

1.3 Deep Learning$_{L}earning_{B}ackground.$

1.4 Thesis Objectives

Chapter 2

BACKGROUND

a u i no-vowel fat-ha, dam-ma, kas-ra and sukun harakat Shadaa Tanween

Definition 1 ***Shadaa***
To indicate the letter is doubled. Any letter with shaddah ( ) should be duplicated: the first letter with a constant (suku

Definition 2 ***Tanween***
This doubles the short vowel, and can convert Tanween fathah, Tanween dhammah or Tanween kasrah by replacing it w

/n/ Sukun noon-Sakinah tanween Tanween noon-Sakinah tanween tanween harakat haraka noon-Sakinah tanween harakaha

preparing data $\{, , , , , , \}e$

2.1 Arabic Arud

Definition 3 ***Arud***
In Arabic, Arud has many meanings (the way, the direction, the light clouds and Mecca and Madinah [1] *[6]. Arud is the*

Al-Farahidi meters Al-Farahidi seas

Al-Farahidi

•
•
•

•*with fat-ha, not with dam-ma such as the rules name*
•

2.1.1 Al-Farahidi and Pattern Recognition

tafa'il

2.1.2 Feet Representation

Definition 4 ***Feet***
A foot consists of a sequence of ***Sukun*** *(Consonants) represented as (0) and* ***Harakah*** *(Vowels) (/). Traditionally, feet*

- **Asbab ()**
  Sabb Khafeef ( )harakah sukun
  Sabb Thakeel ( )harakah
- **Awtad ()**
  Watd Magmo'a ( )harakah sukun
  Watd Mafrouq ( )harakah sukun
- **Fawasek (**
  Faselah Soghra (
  Faselah Kobra ([2]

2.1.2.1 Rules for Arabic Letter Representation

2.2 Deep Learning BackgroundWhat is Deep Learning? *Deep Learning is a new approach to Machine Learnin*

Naive BayesLogistic regression

2.2.1 Logistic Regression

Sigmoid function$_f$*unction.TheLogisticfunctionisshapedasanS* $-$ *shape.*

$\widehat{y}y\widehat{y}_e stimate.So, to calculate the output function for Logistic Regression using Equation eq : logistic_r egression_y hat.No$

2.2.1.1 Loss Error Function$\widehat{y}y\widehat{y}^5{}_f unction describe the loss function for Logistic Regression. There are other functions$

$_f unction we need \widehat{y}y$ *true*$-(\log \widehat{y})$
$_f unction we need \widehat{y}y$ *true*$-\log(1-\widehat{y})$

2.2.1.2 Cost Function$y\widehat{y}$***(w,b)***$_r egression_y hat as$ ***(w,b)***$are the parameters which define the relation between input data s$

$J_f unction is the average of loss function applied to every training example which equals the sum of the loss for each trainin$
$= \frac{-\sum_{i=1}^{m}[(y^i \log \widehat{y^i} + (1-y^i)\log(1-\widehat{y^i}))]}{m}$ (2.5)

2.2.1.3 Convex Function vs Non-Convex Function

Convex Function$convex f(x) convex [a,b] x_1 x_2 [a,b] \lambda 0 < \lambda < 1$

$convex X^2 f(x) [a,b] convex f''(x) \geq 0 x [a,b]$
$x_1 x_2 f(x)$**strictly convex***Convex function*
Non-Convex Function*convex function*$f(x)[a,b]x_1x_2[a,b] - f(x)$

$fff'' f(x) = x^4 ff`$

2.2.1.4 Gradient Descent***(w,b)***$_f unction to the minimum. In other words we need to find the best value of$ ***J(w,b)***$which w$

$f(x) = x^2 P_1 P_2$*(which by definition is the slope of the function at the point which also the change between these two*

(w,b)J(w,b)$_d escent_{jw} wrt$ ***(w)***$, and second function in Equation eq : gradient_d escent_{jb} wrt$ ***(b)***$w := w - \alpha dw alpha$ *is b*
$:= w - \alpha\frac{dJ(w)}{dw}$ *d represent the derivative wrt* w(2.9)$\alpha\frac{dJ(w,b)}{dw}$
$b := b - \alpha\frac{dJ(w,b)}{db}$

2.2.1.5 Logistic Regression derivatives$\widehat{y}_r egression_d erivatives_s ingle_e xample. So, doing backpropagation to get the va$

Chapter 3

Literature Review

3.1 Deterministic (Algorithmic) Approach *learning problem* *deterministic five-step algorithm* *Arud writing* *if-else* *A*

if-else *regular expressions* *harakat*

The size of the test data
The step converting verses into ones and zeros pattern

3.2 English Literature

Chapter 4

DESIGN DATASET

1

4.1 Dataset Design

- **Datasets availability:**
- **The Poem with diacritics:**
- **The amount of the dataset:**
- **Cleansing of this data:**

- **Datasets availability:**
- **The size of the dataset:**
- **Poems with diacritics:**
  -
  -
  -
  -
- **Cleansing of this data:** $_Clens.Wealsoopensourcedallthecodescriptsusedinouronlinerepository$[].

4.1.1 Data Scraping

4.1.2 Data Preparation and Cleansing

-
-
-
-
-
-
- *harakah*
-
- [1]

4.1.3 Data Encoding

4.1.3.1 Encoding in English

- **Word embedding Encoding in English:**
- **Character Level Encoding in English**

4.1.3.2 Character Level Encoding in Arabic *one-hot binary one-hot binary*

binary one-hot

$nn \times pnpn$ *one-hot binary* ***two-hot***

one-hot $37 \times 137 \times 5$ *one-hot* 41×9

one-hot $37 \times 137 \times 5$ *one-hot* 41×9

RNN RNN

one-hotbinarytwo-hot

JJ

+=

4×1

37×1

One-Hot Binary Two-Hot

***One-hot* encoding** $36 \times 5181 \times 1$ *one-hot*
one-hot one-hot $181 \times 1181 \times 82$
Binary Encoding $n \times 1 n = * \log_2 l l n l$ *binary* 8×5 *one-hot binary*
***two-hot* encoding** *one-hot* 37×1 *one-hot* 4×1 *one-hot* 41×1

4.1.3.3 Categorical Values Encoding *label encoding, one hot encoder, and custom binary encoding one-hot Sklearn*

Chapter 5

Model Training

• $_{R}ep_{P}aram.Datarepresentationfeatureisaffectedby Arabiclanguagepronunciation, andsomefeaturesprovidemoreinfor$
• $_{P}aram.The number of verses (344, 464) usedintesting and validation is significant, confirming that the model wastested on a$

5.1 Parameters of Data Representation

5.1.1 Diacritics$_{E}ncoding, the inclusion or otherwise of diacritics has the same length in input vector size.$

5.1.2 Trimming Small Classes$_{L}oss. The trimmed classes are five classes which have less than 1\% of the total dataset. W$

5.1.3 Encoding Techniques$_{E}ncoding. Although all carry the same information, it was expected that every encoding has i$

- **Running Time**
- **Required Resources**
- **Learning Rate**
- **Overall Performance**
- **Overall Performance**

5.1.4 Data Representation Matrix*one-hot*

5.2 Parameters of Network Configuration$_{L}stm, as an alternative way to test the effect of BI - Directional LSTM,$

tafa'il

- **Cell Type**
- **Layers**
- **Cell Unit Size**
- **Weighting Model**$_{L}oss to help work on all the dataset. We therefore have two combinations; one with weighting loss and one w$

$422^4 = 1616 \times 12 = 19296 \times 2 = 192\{Trimming, NoTrimming\} \times \{Diacritics, NoDiacritics\} \times \{OneE, BinE, TwoE$

5.2.1 Working on Unbalanced data using Weighted Loss

1

$$c = \frac{\frac{1}{288}}{\sum \frac{1}{416428} + \frac{1}{370116} \cdots + \frac{1}{288}}$$

$$= \frac{\frac{1}{288}}{0.00535} = 0.03$$

$$= \frac{\frac{1}{416428}}{0.00535} = 0.0004$$

5.3 Experiments

5.4 Hardware[2]

[3]

5.5 Software

- *Used as main programming language.*
- *Used as Deep learning backend framework*
- *Used as High level framework on top of the backend*
- *Used in data pre-processing and cleansing.*
- *Used in data pre-processing and splitting.*
- *Used to encode the classes using Label-Encoder and for model assessment phase.*
- *Used to save the encoder and the model as serialized pickle object.*
- *Used to save the encoded dataset matrix in h5 format.*

5.6 Implementation Outline

- *Numpy*
- [4]
-
- [5]
-
- *Shadaa and Tanween*
-
- *Sklearn*
-
- *PickleH5 formath5py*
- *one-hotbinarytwo-hot*
- *Full/Eliminated and With/Without tashkeeltashkeel*
- *h5*
- *Keras with Tensorflow*
- *Full/Eliminated and With/Without tashkeelLayers, Units and cell type (Bi-LSTM or LSTM)*
-
- –
- –
- –
-
-
- *Tensorflow*
-

[draw,fill=white] at (4.2,-1.75) [/pgfplots/every crossref picture]/pgfplots/legend image code[,/tikz/.cd,bar width=1pt,y

Chapter 6

Results

Text
26
16

15

(14)
20
6

6.5 Encoding Effect$_{E}ncoding. In this section, we will explore the effects of Data Encoding with respect to Accuracy, Lea$

two-hot *binary* *one-hot*

one-hot $181 \times 8(bits) 1,448$ *two-hot* $41 \times 8(bits) 328$ *binary* $8 \times 8(bits) 64$ *two-hot*$_{E}ncoding$

6.6 Comparison with Literature

6.7 Classifying Arabic Non-Poem Text

tashkeel *tashkeel* *Al-Taweel*

Harakat and Sukun *Al-Taweel* *Tafail*

tafa'il *harakah* *harakah* *Al-Taweel* *tafa'il* *Al-Taweel* *tafa'il*

Example:

```
0/0//0/0/0///0//0//0//0/0//0/0/0//OliveGreen/0//red/OliveGreen00//0//
0/0//0/0/0// /0//0/0///0/0//0/0/0//0/0//0//0//
```

Tafail

tashkeel *tashkeel* *tashkeel* *tafa'il* *Al-Taweel* *tashkeel* *tafa'il* *tashkeel* *tafa'il*

Example:

```
0/0//0/0/0////red /0/0/0/0///red 0/00/0/0/0////red /0///red 0/0/
0/0//0/0/0// 0/0//0/0///0/0//0/0/0//0/0//0//0//
```

6.8 Discussion

6.8.1 Dataset Unbalanced

6.8.2 Encoding Method

$\mathcal{T} X \mathcal{T}(X) \mathcal{T}(X) \eta_1 \left(\mathcal{T}_1(X)\right) = \left(\eta_1 \cdot \mathcal{T}_1 \cdot \mathcal{T}_2^{-1}\right) \left(\mathcal{T}_2(X)\right) \eta_1 \mathcal{T}_1 \eta_1 \mathcal{T}_2 \mathcal{T}_2 \eta_2 = \eta_1 \cdot \mathcal{T}_1 \cdot \mathcal{T}_2^{-1} \mathcal{T}_2(X)$

$\mathcal{T} \eta$ *two-hot one-hot binary*

6.8.3 Weighting Loss Function

6.8.4 Neural Network Configurations

6.8.5 Model Assessment F_1

Chapter 7

Conclusion and Future Work

7.1 Conclusion*diacritics*

7.2 Future Work

-
  - \{}
-

Bibliography

Fundamentals of Deep Learning: Designing Next-generation Machine Intelligence Algorithms
JSW
Deep Learning
Principles of Mathematical Analysis
The Secret Life of the Brain
CoRR
Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 20
IEEE Transactions on Neural Networks
CoRR
Neural Comput.
CoRR
CoRR
$_config_en.html$
International Arab Journal of Information Technology
J. of Commun. & Comput. Eng.
Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers
FLAIRS Conference
Proceedings of the Fifth Workshop on Computational Linguistics for Literature
Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Ranlp
J. Mach. Learn. Res.
J. Mach. Learn. Res.
Advances in Neural Information Processing Systems 26
Empirical Methods in Natural Language Processing (EMNLP)
CoRR
CoRR
Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17
Journal of Machine Learning Research