

PENGECAMAN HURUF ARAB MENGGUNAKAN TEKNIK RANGKAIAN NEURAL

NUR SURIZA SYAZWANY AHMAD NIZAM
KHAIRUDDIN OMAR

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

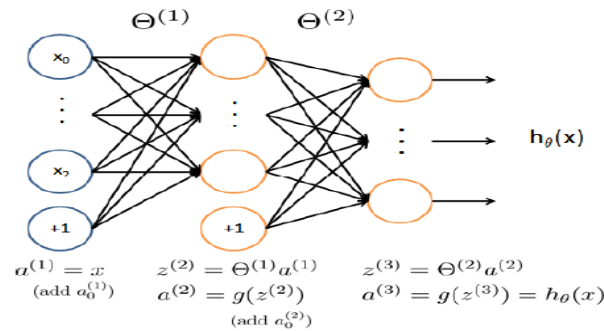
Dalam kajian projek ini, teknik rangkaian neural digunakan bagi mengkaji pengecaman huruf Arab. Kajian projek ini menggunakan algoritma perambatbalikan. Pengubahsuaian dilakukan terhadap niali fungsi kos (*cost function*), kadar pembelajaran (*learning rate*) dan saiz lapisan tersembunyi (*hidden layer sizes*) bagi melihat kesan terhadap kadar ketepatan pengecaman. Proses pengecaman ini dilaksanakan terhadap fungsi kos logistic dan fungsi kos logistic secara rombongan. Kedua-dua fungsi kos ini dilarikan mengikut 3 jenis pengoptimum iaitu *Quasi-Newton(lbfgs)*, *stochastic gradient descent(sgd)* dan *adaptive moment estimation (adam)*. Hasil daripada pengujian yang dijalankan, kadar ketepatan meningkat apabila kadar pembelajaran (*learning rate*) meningkat. Pengoptimum jenis *Quasi-Newton* merekod kadar ketepatan tertinggi berbanding *Adaptive Moment Estimation (adam)* dan *Stochastic Gradient Descent (sgd)*.

1 PENGENALAN

Huruf Arab mempunyai 28 huruf dan pada dasarnya kelihatan berkeluk. Huruf Arab ditulis secara mendatar bermula dari kanan ke kiri. Huruf Arab adalah unik dan sangat berbeza dengan huruf lain seperti huruf latin dan huruf cina. Huruf Arab tidak tetap dan berubah-ubah saiz, orientasi dan kedudukan. Kebanyakan huruf Arab mempunyai 4 jenis bentuk: asing, awal, tengah dan akhir. Bentuk huruf bergantung kepada kedudukan huruf tersebut samada ianya huruf tunggal, awal perkataan, tengah perkataan atau di akhir perkataan.

Banyak kajian telah dilaksanakan berkaitan dengan pengecaman tulisan tangan termasuklah pengecaman huruf tulisan tangan. Pelbagai teknik dan model yang berbeza digunakan untuk membuat pengecaman termasuklah teknik Hidden Markov Model dan Rangkaian Neural. Teknik Rangkaian Neural akan diaplikasikan ke atas 616 data (Dr. Husam) yang terdiri daripada pelbagai jenis bentuk bagi setiap satu huruf Arab.

Rajah 1.1 menunjukkan bentuk Rangkaian Neural terdiri daripada 3 lapisan iaitu lapisan pemasukan (*input layer*), lapisan tersembunyi (*hidden layer*) dan lapisan pengeluaran (*output layer*).



Rajah 1.1 Model Rangkaian Neural

Sekiranya imej bersaiz 20x20, ini bermakna 400 lapisan pemasukan (*input layer*) akan diperlukan.

2 PENYATAAN MASALAH

Setiap satu huruf Arab mempunyai pelbagai bentuk berlainan. Huruf Arab tidak tetap dan berubah-ubah saiz, orientasi dan kedudukan. Bentuk huruf dibahagikan kepada 4 iaitu huruf tunggal, awal perkataan, tengah perkataan dan akhir perkataan.

Huruf Arab juga mirip antara satu sama lain dimana boleh mengakibatkan kekeliruan. Kebanyakan huruf hanya dibezakan dengan jumlah titik yang berbeza antara satu huruf dengan huruf yang lain.

Tulisan tangan huruf Arab tidak terlalu jelas serta kabur mengakibatkan kesukaran untuk mengenali huruf tersebut.

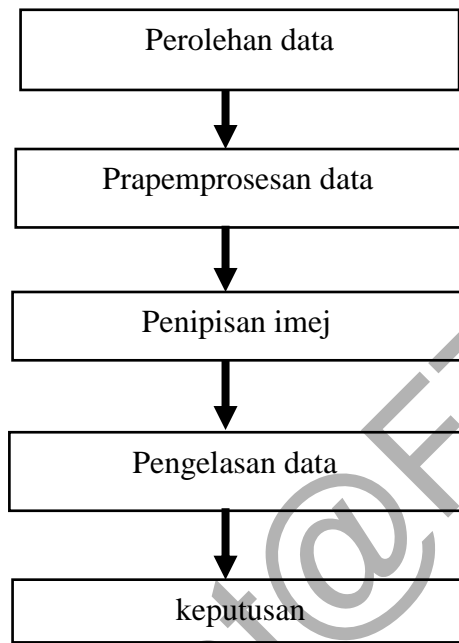
3 OBJEKTIF KAJIAN

Mengkaji kesan terhadap kadar ketepatan pengecaman huruf Arab sekiranya pengubahsuaian dilakukan ke atas faktor-faktor tersebut:

- I. Fungsi kos (Cost function)
- II. Kadar pembelajaran (learning rate)
- III. Bilangan nod lapisan tersembunyi (hidden layer nodes)

4 METODOLOGI KAJIAN

Dalam projek pengecaman ini, metodologi adalah penting supaya proses pengecaman berjalan lancar. Rajah 3.1 menunjukkan metodologi yang akan digunakan dalam proses pengecaman huruf Arab menggunakan teknik rangkaian neural ini.



Rajah 4.1 Metodologi Kajian

4.1 PEROLEHAN DATA

Data sebanyak 616 imej huruf Arab tunggal bertulisan tangan pelbagai saiz diperoleh melalui dataset Dr. Husam. Huruf Arab tunggal yang dimaksudkan terdiri daripada bentuk iaitu asing, awal, tengah dan akhir.

4.2 PRAPEMPROSESAN DATA

Setiap data yang diperoleh daripada dataset Dr. Husam terdiri daripada pelbagai saiz imej. Setiap imej tersebut akan diselaraskan kepada 20 piksel x 20 piksel dan seterusnya menukar format fail kepada fail bitmap (.bmp) untuk memudahkan proses pengecaman dijalankan. Setiap imej juga turut ditukarkan kepada skala kelabu (grayscale).

4.3 PENIPISAN IMEJ

Setelah imej diselaraskan kepada 20x20 piksel, imej-imej tersebut akan dijadikan 400-dimensi vektor. Setiap data latihan menjadi satu barisan dalam data matriks X. Oleh itu, sebanyak 616 data dijadikan 400 matriks X setiap satunya dimana setiap baris merupakan contoh latihan bagi imej huruf Arab tulisan tangan.

Bahagian seterusnya merupakan 616-dimensi vector y yang mempunyai label untuk set latihan. Pengindeksan dilakukan daripada nilai sifar sehingga nilai 20 memandangkan hanya terdapat 20 huruf Arab dalam dataset ini. Oleh itu, nilai '1' dilabel sebagai huruf Alif dan seterusnya.

4.4 PENGELASAN DATA

Rajah 4.2 menerangkan proses pengelasan data menggunakan teknik Rangkaian Neural melalui carta aliran.

I. Memuatkan data

Data imej bersaiz 20 piksel x20 piksel dimuatkan.

II. Memuatkan parameter

Pemboleh akan dimuatkan ke dalam pembolehubah Theta1 dan Theta2. Kemudian parameter akan dibuka.

III. Pengiraan fungsi kos

Pengiraan fungsi kos dilaksanakan secara tanpa rombakan dimana :

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K \left[-y_k^{(i)} \log((h_{\theta}(x^{(i)}))_k) - (1 - y_k^{(i)}) \log(1 - (h_{\theta}(x^{(i)}))_k) \right]$$

IV. Melaksanakan rombakan

Pengiraan fungsi kos dilaksanakan secara rombakan dimana:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K \left[-y_k^{(i)} \log((h_{\theta}(x^{(i)}))_k) - (1 - y_k^{(i)}) \log(1 - (h_{\theta}(x^{(i)}))_k) \right] + \frac{\lambda}{2m} \left[\sum_{j=1}^{25} \sum_{k=1}^{400} (\Theta_{j,k}^{(1)})^2 + \sum_{j=1}^{10} \sum_{k=1}^{25} (\Theta_{j,k}^{(2)})^2 \right].$$

V. Melaksanakan kecerunan sigmoid

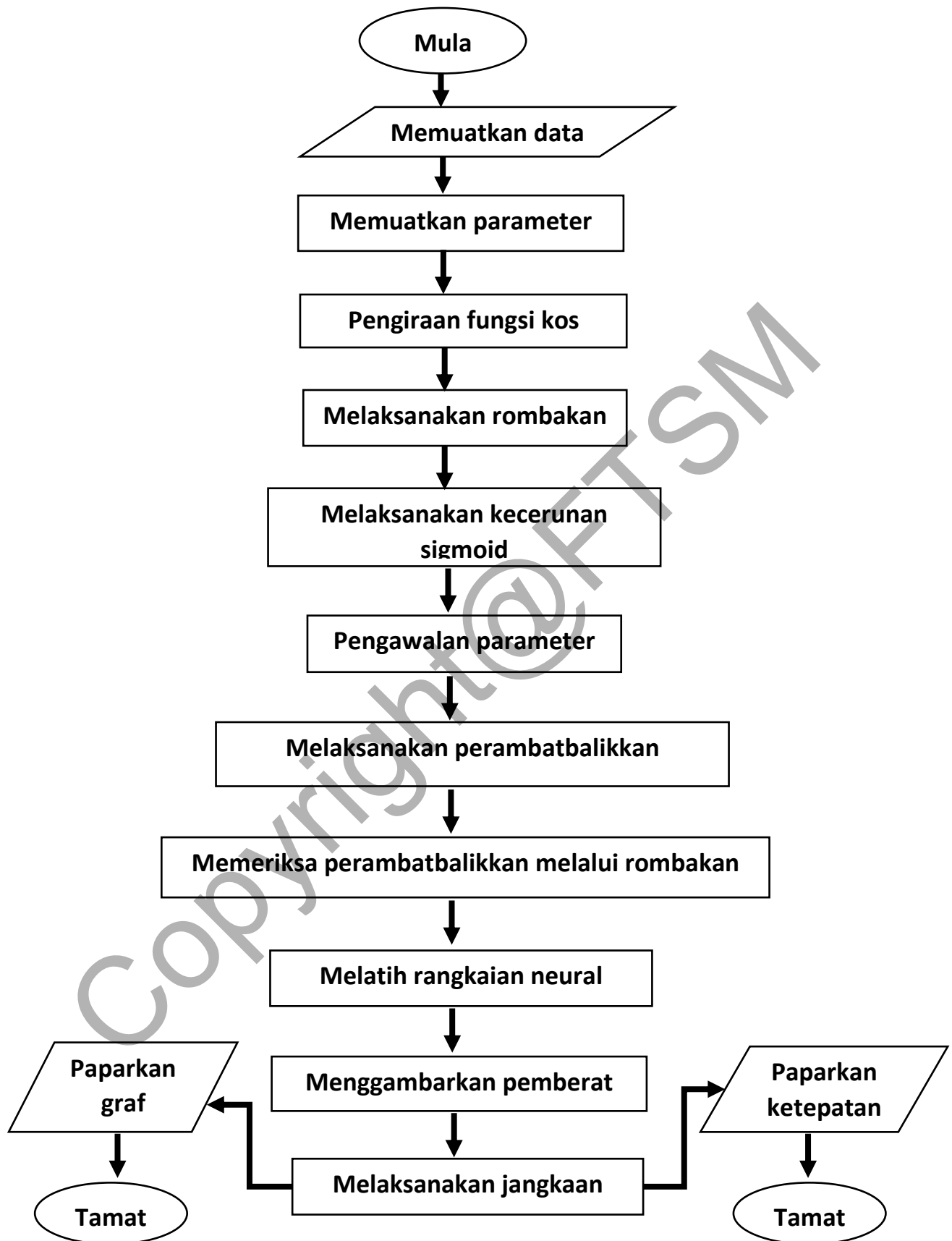
Melaksanakan kecerunan terhadap fungsi sigmoid dengan

$$g'(z) = \frac{d}{dz}g(z) = g(z)(1 - g(z))$$

dengan

$$\text{sigmoid}(z) = g(z) = \frac{1}{1 + e^{-z}}.$$

Copyright@FTSM



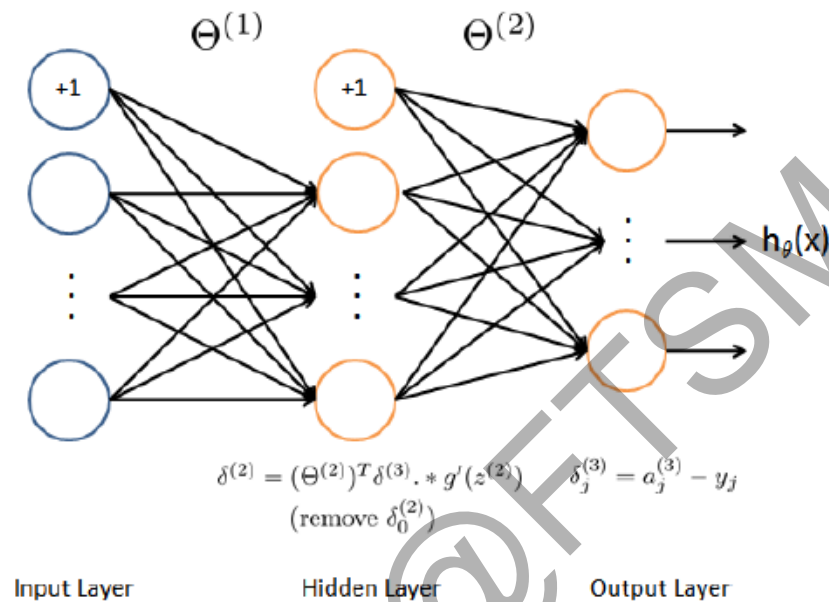
Rajah 4.2 Carta aliran proses pengelasan data menggunakan teknik Rangkaian Neural

VI. Pengawalan parameter

Mulakan dua lapisan rangkaian neural dengan memulakan pemberat dan parameter.

VII. Melaksanakan perambatbalikkan

Rajah 4.3 menunjukkan pelaksanaan perambatbalikkan.



Rajah 4.3 pelaksanaan perambatbalikkan

VIII. Memeriksa perambatbalikkan dengan melakukan rombakan

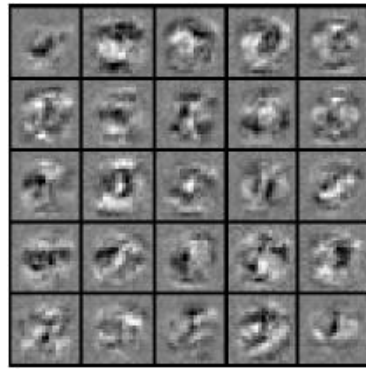
Rombakan dilaksanakan ke atas perambatbalikkan agar fungsi kos dan kecerunan dapat dilakukan.

IX. Melatih rangkaian neural

Rangkaian neural akan dilatih menggunakan pengoptimal lanjutan setelah dilaksanakan pengiraan kecerunan.

X. Menggambarkan pemberat

Antara cara untuk memudahkan pemahaman terhadap proses rangkaian neural yang berlaku adalah dengan menggambarkan unit tersembunyi untuk melihat ciri-ciri tangkapan data. Rajah 4.4 menunjukkan gambaran lapisan tersembunyi.



Rajah 4.4 Gambaran lapisan tersembunyi

XI. Melaksanakan jangkauan

Fungsi jangkauan dilaksanakan untuk meramal label set data. Ia juga membolehkan pengiraan ketepatan pengecaman dilaksanakan.

5 HASIL KAJIAN

Kadar ketepatan dikeluarkan untuk melihat keberkesanan data latihan. Fungsi kos atau fungsi kehilangan turut dikeluarkan dan dijadikan graf bagi melihat perbezaan keberkesanan ketiga-tiga jenis pengoptimum tersebut. Sebanyak 4 set ujikaji dijalankan menggunakan kadar pembelajaran pengawalan (*learning_rate_init*) yang berbeza iaitu 0.1, 0.01, 0.001, 0.0001.

Jadual 5.1. Kadar ketepatan fungsi logistic

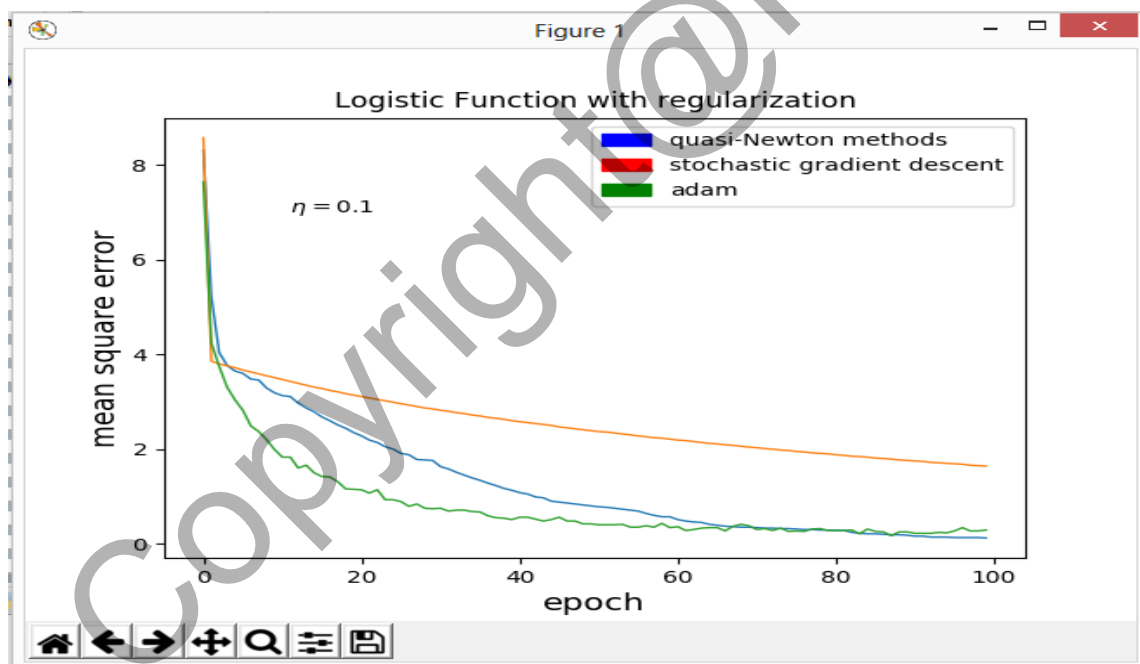
FUNGSI LOGISTIC				
η	0.1	0.01	0.001	0.0001
Quasi-Newton	21.6%	21.6%	21.6%	21.6%
Stochastic Gradient Descent	10.8%	0	5.4%	5.4%
Adam	21.6%	24.3%	10.8%	10.8%

Jadual 5.2. Kadar ketepatan fungsi logistic with regularization

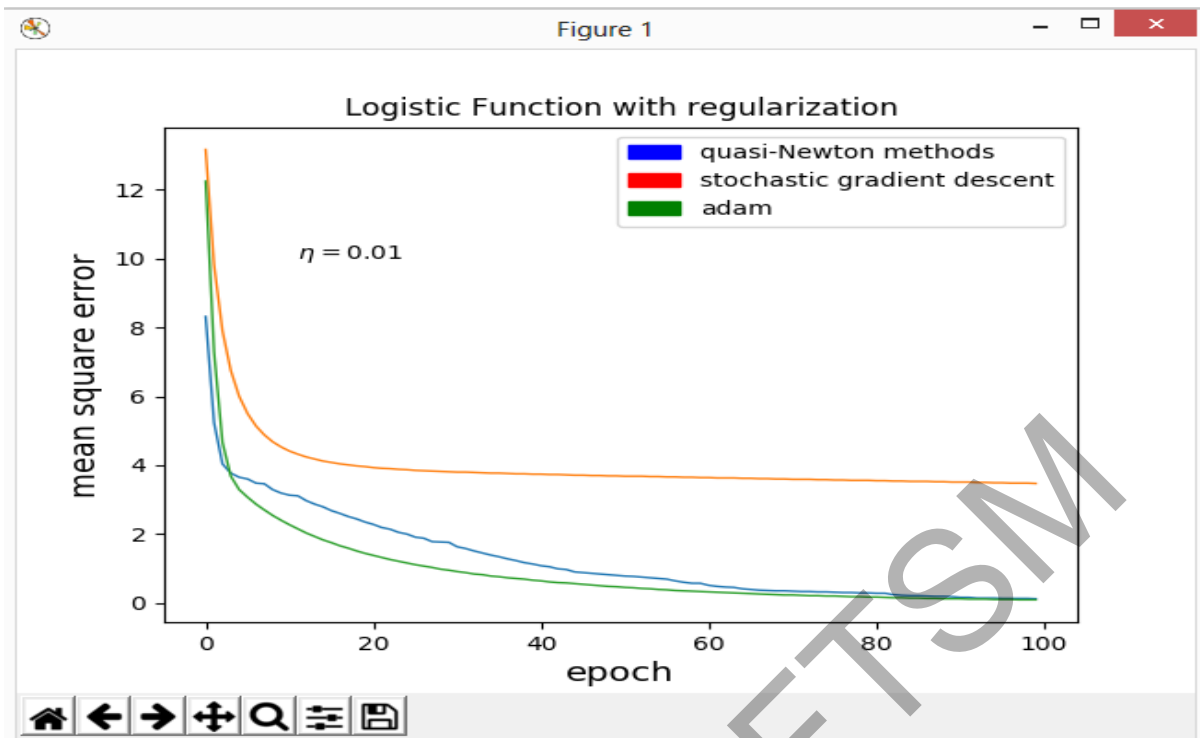
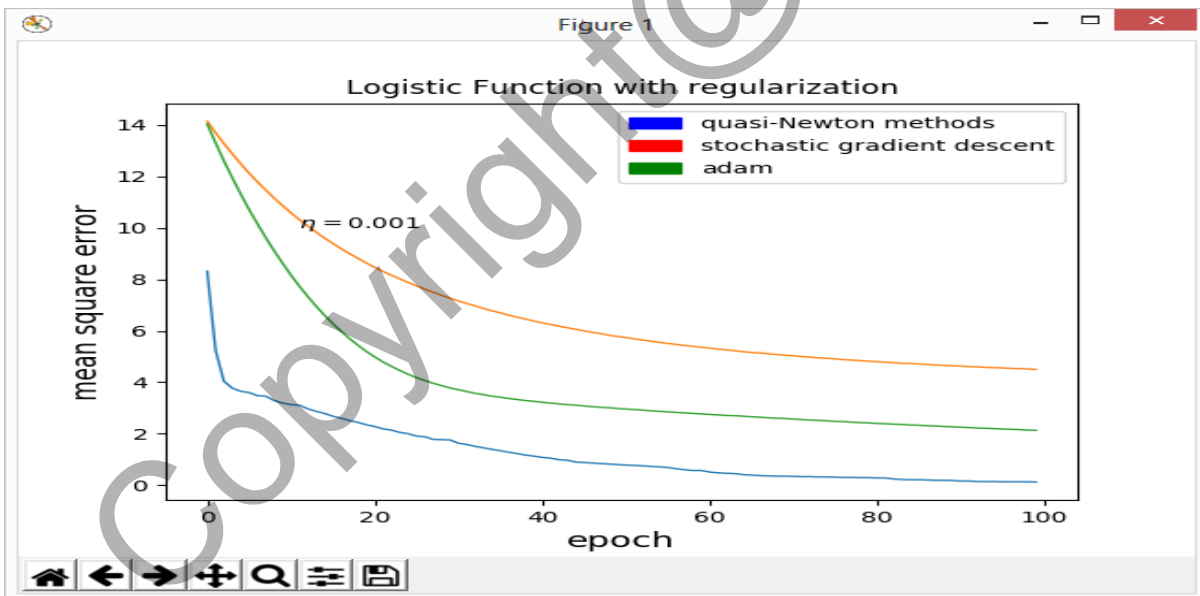
FUNGSI LOGISTIC WITH REGULARIZATION				
η	0.1	0.01	0.001	0.0001
α	0.01	0.01	0.01	0.01
Quasi-Newton	21.6%	21.6%	21.6%	21.6%
Stochastic Gradient Descent	10.8%	0	5.4%	5.4%
Adam	29.7%	18.9%	10.8%	13.5%

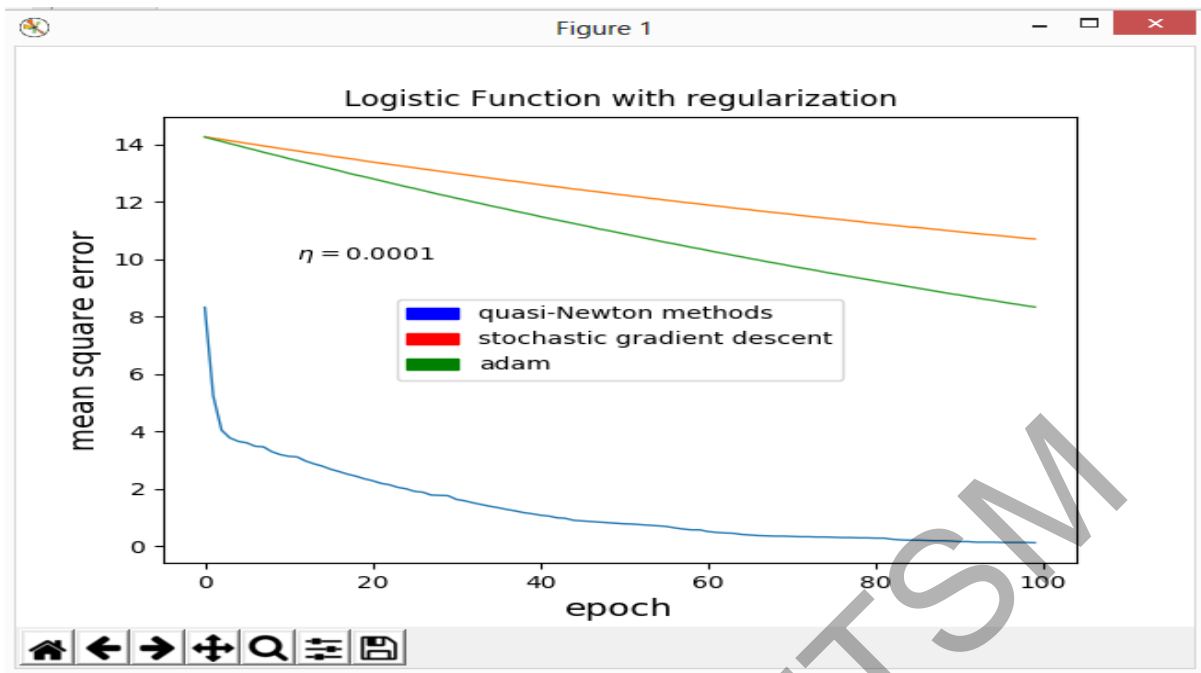
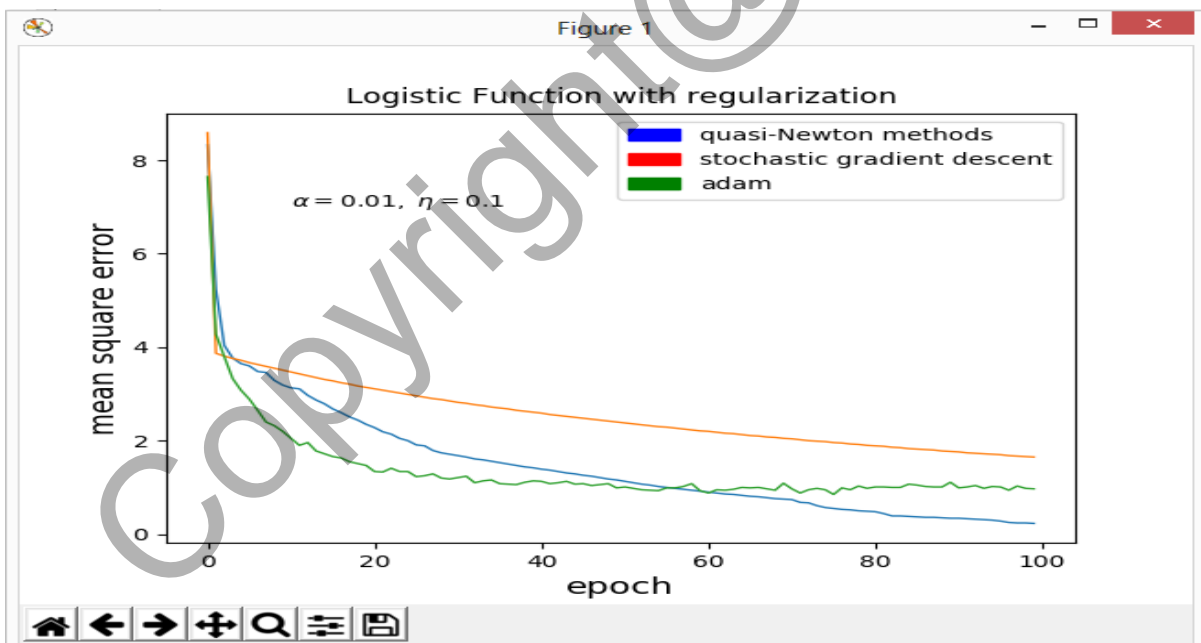
Kajian mendapati kadar ketepatan pengecaman huruf Arab dipengaruhi oleh nilai kadar pembelajaran iaitu semakin tinggi kadar pembelajaran, semakin tinggi kadar ketepatan pengecaman seperti yang dipaparkan di dalam hasil Graf fungsi *logistic* ($\eta=0.1$) (Rajah 5.1), Graf fungsi *logistic* ($\eta=0.01$) (Rajah 5.2), Graf fungsi *logistic* ($\eta=0.001$) (Rajah 5.3), dan Graf fungsi *logistic* ($\eta=0.0001$) (Rajah 5.4). Manakala Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.1$) (Rajah 5.5), Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.01$) (Rajah 5.6), Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.001$) (Rajah 5.7), dan Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.0001$) (Rajah 5.8).

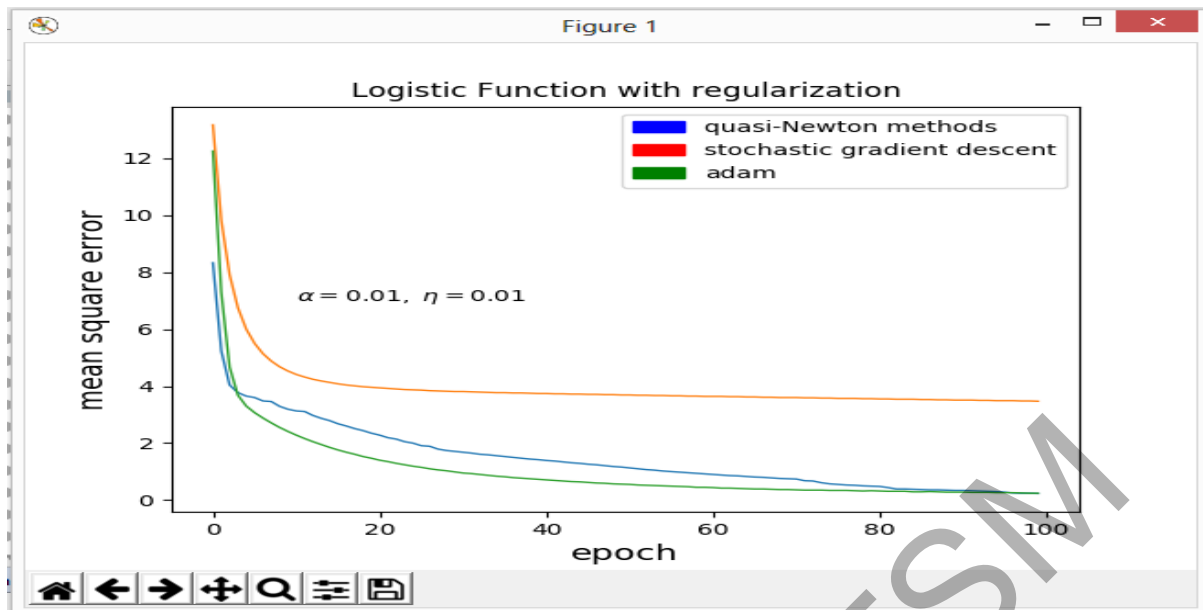
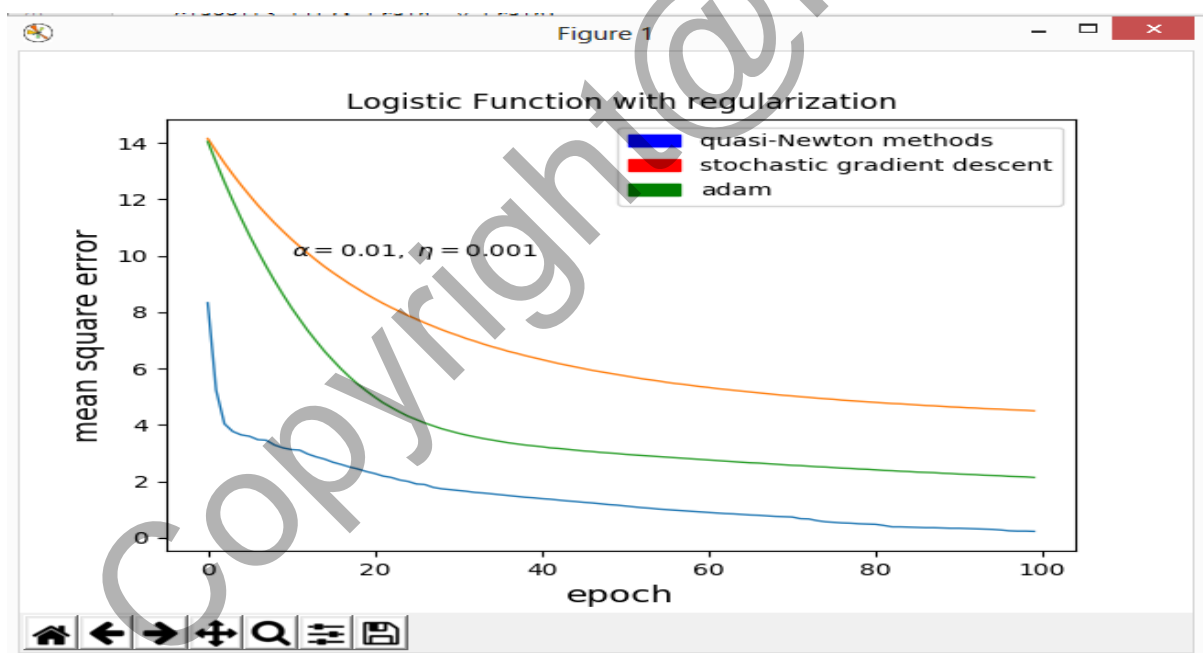
Kajian juga mendapati pengoptimum jenis Adam merekod kadar ketepatan paling tinggi pada nilai kadar pembelajaran 0.1 setelah melaksanakan rombakan, iaitu sebanyak 29.7% manakala kadar ketepatan terhadap jenis pengoptimum *Stochastic Gradient Descent* (*sgd*) pada nilai pembelajaran 0.01 mencatat kadar ketepatan paling rendah iaitu 0%.

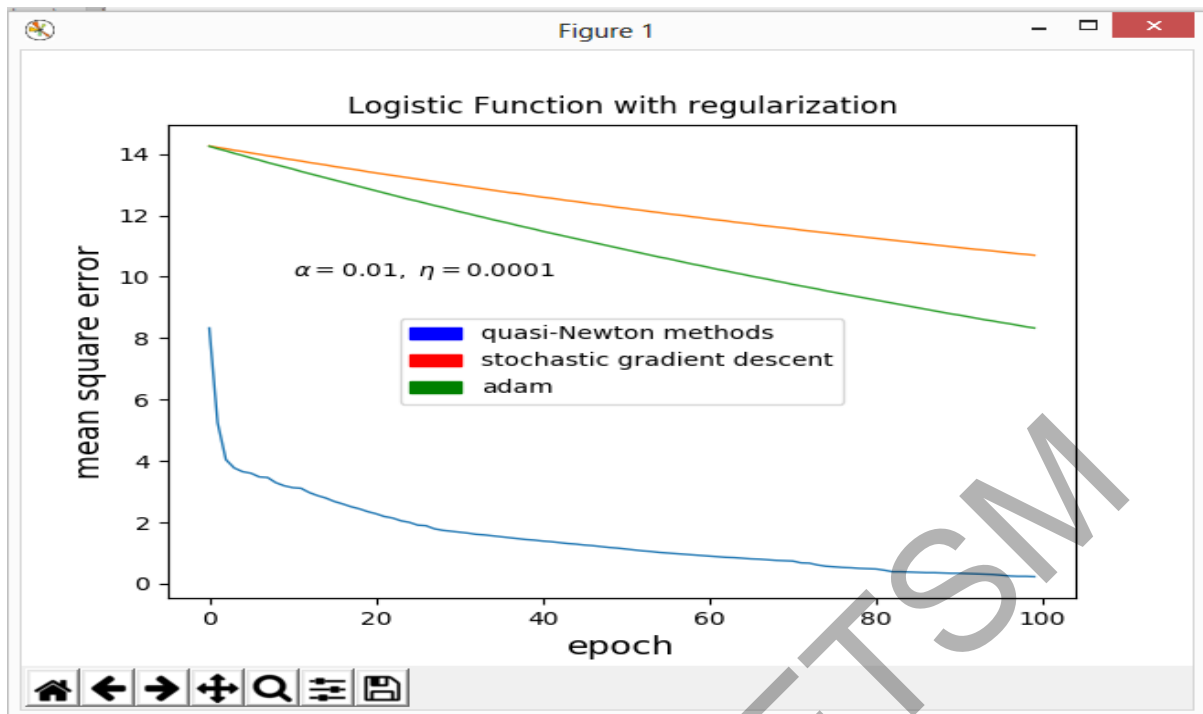


Rajah 5.1. Graf fungsi *logistic* ($\eta=0.1$)

Rajah 5.2. Graf fungsi *logistic* ($\eta=0.01$)Rajah 5.3. Graf fungsi *logistic* ($\eta=0.001$)

Rajah 5.4. Graf fungsi *logistic* ($\eta=0.0001$)Rajah 5.5. Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.1$)

Rajah 5.6. Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.01$)Rajah 5.7. Graf fungsi *logistic with regularization* ($\alpha=0.01, \eta=0.001$)



Rajah 5.8. Graf fungsi *logistic with regularization* ($\alpha=0.01$, $\eta=0.0001$)

6 KESIMPULAN

Dalam kajian ini, proses pengecaman huruf Arab menggunakan teknik rangkaian neural telah berjaya mencapai objektif kajian di mana pengubahsuaian yang dilakukan ke atas fungsi kos (*cost function*), kadar pembelajaran (*learning rate*), bilangan nod lapisan tersembunyi (*hidden layer*) didapati memberi kesan terhadap kadar ketepatan pengecaman. Fungsi kos rombakan (*regularized cost function*) didapati mempunyai tahap ketepatan lebih tinggi daripada fungsi kos tanpa rombakan (*non-regularized cost function*). Kadar ketepatan pengecaman juga didapati meningkat apabila nilai kadar pembelajaran (*learning rate*) meningkat.

Dari segi jenis pengotimum pula, *Quasi-Newton* mempunyai kadar ketepatan tertinggi diikuti dengan *Adaptive Moment Estimation (adam)* dan akhir sekali *Stochastic Gradient Descent (sgd)*.

Kelebihan kajian pengecaman huruf Arab menggunakan teknik rangkaian neural secara perambatbalikan ini ialah:

- I. Kajian ini mampu untuk membuat pengecaman dalam masa yang singkat.

- II. Kajian ini hanya memerlukan data bersaiz kecil sekaligus menjimatkan ruang simpanan data.

Antara kekurangan kajian pengecaman huruf Arab menggunakan teknik rangkaian neural secara perambatbalik ialah:

- I. Kadar ketepatan pengecaman yang rendah
- II. Data yang banyak diperlukan untuk meningkat kadar ketepatan pengecaman

7 RUJUKAN

- Ipin Supriana & Albadr Nasution, 2013. Arabic character recognition system development in The 4th International Conference on Electrical Engineering and Informatics (ICEEI 2013), pp. 34-341.
- Chirag I Patel, Ripal Patel & Palak Patel, 2011. Handwritten Character Recognition using Neural Network, International Journal of Scientific & Engineering Volume 2, Issue 5, pp. 1-6.
- Jamril Ramdan 2014. Developing An Arabic Handwriting Database for Text Recognition (AHDB-FTR), The 3rd National Doctoral Seminar on Artificial Intelligence Technology, pp. 165-170.