

shirey@ce.aut.ac.ir

dezhkam@ce.aut.ac.ir

shahmiri@ce.aut.ac.ir

[]

[]

%

[]

%

[]

% /

" "

" "

[]

" "

⁶ Text Categorization
⁷ Knowledge-based

¹ Poet Identification
² Author Identification
³ Natural Language Processing
⁴ Feed-Forward Backpropagation
⁵ Single Perceptron

WTMM¹⁴ DFA¹⁵ " " []

" " []

% / % / " " []

[]

" " []

[]

" "

" % " " []

" " []

[]

" "

k

% [] %

" "

" " []

[]

¹⁴ Detrended Fluctuation Analysis
¹⁵ Wavelet Transform Modulus Maxima
¹⁶ Fluctuation Analysis of Sequences
¹⁷ Light Support Vector Machine (SVM^{light})
¹⁸ Feature Reduction

⁸ Augmented Relevancy Signatures Algorithm
⁹ Term Selection
¹⁰ Latent Semantic Indexing (LSI)
¹¹ Hierarchically Self-Organizing Maps (HSOM)
¹² Letter Sequence
¹³ k -Nearest Neighbour

[]

[]

)

[] () (

()

[] ()

(

...	/	/	سجل مرسوم
...	.../		سجل
...			سجل
...	.../	/	سجل (مستوفى)

() : " / "

() : " / "

() () " " " "

[]

() () " "

[]

" " ()
 " " ()
 " " ()
 " " ()
 " " ()

"-U--U--U--U" ()

"-" "U"
 "v" "c"

"V"

[]

(

U	cv
-	cV
-	cvc
-U/-	cVc
-U/-	cVcc / cVcc

[]

[]

" " : []

" " []

[]

[]

" "

()

()

" " ()

manamgoftbAfarreyezadi
 hamamSahreyArihamammawbadi

مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب
 مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب مذہب

[]

[]

cvcvcvcccVcvcvcvcVcvcV
 cvcvcvcvcVcVcvcvcvcvcV

()

" " " " " " " " " " " "

$$y = f\left(\sum_{i=1}^n (w_i x_i + w_0)\right)$$

"kahon" "saxon" " " " "

()

[]

)

w x
[]

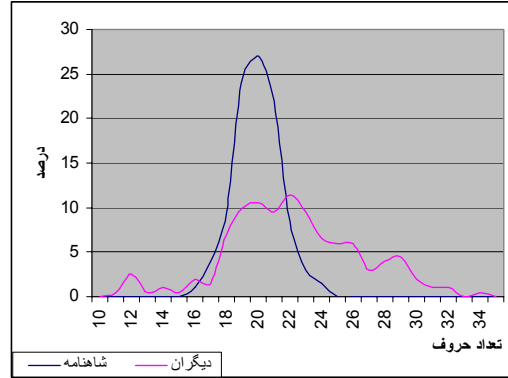
y

(

()

$$\alpha = \sum_{i=0}^n a_i x_i / n$$

a_i x_i n α

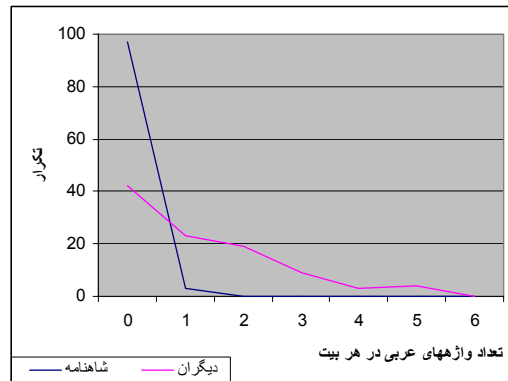


()

)

()

(



()

(III) (II) (I)

()

/	/	/	/	I
			/	II
				III
/	/	/	/	II I
				III I
/	/	/		III II
/	/		/	III II I

∞

)

(

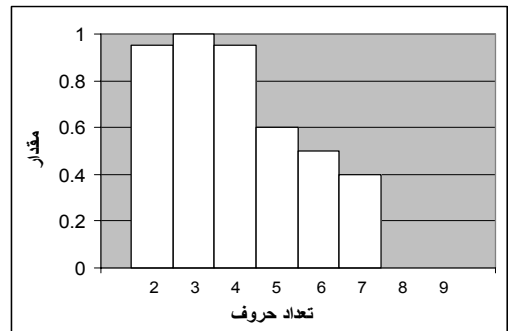
[]

(

)

(

)



()

(I)	(VII IV)	(III)	(II)
()	() ()	()	
	/	/	I
	/	/	II
	/	/	III
	/		II I
	/		III I
	/		III II
	/		III II I
	/		IV
	/	/	V
	/	/	VI
		/	VII

VII IV

)
(
()

%

:

%

%

% /

:

- [1] Russell S. J., And Norvig P. *Artificial Intelligence: A Modern Approach*. Prentice-Hall International Inc., 1994.
- [2] James A. *Natural Language Understanding*. The Benjamin/Cummings Publishing Co., 2nd Edition, 1994.
- [3] Tanimoto S. L. *The Elements of Artificial Intelligence Using Lisp*. Computer Science Press, 2nd Edition, 1995.

[16] Pavlov A. N., Ebeling W., Molgedey L., Ziganshin A. R., And Anishchenko V. S. *Scaling features of texts, images and time series*. Physica A 300, pp. 310–324, 2001.

[17] De Vel O., Corney M., Anderson A., and George Mohay. *Language and Gender Author Cohort Analysis of E-mail for Computer Forensics*, Queensland University of Technology, 2003.

[18] Fukunaga K. *Statistical Pattern Recognition*. Academic Press Inc, 1990.

[19] Webb A. R. *Statistical Pattern Recognition*. John Wiley & Sons Ltd., 2nd Edition, 2002.

[20] Sonka M., Hlavac V., and Boyle R. *Image Processing, Analysis and Machine Vision*. Chapman & Hall Computing, 1993.

[21] Hetch-Nielsen R. *Neurocomputing*. Addison-Wesley Publishing Company, 1989.

[22] Mitchel T. M., *Machine Learning*. McGraw-Hill, 1997.

[4] Maron M. E. *Automatic indexing: an experimental inquiry*. Journal of the Association for Computing Machinery, Vol. 8, no. 3, pp. 404-417, 1961.

[5] Li W., Lee B., Krausz F., and Sahin K. *Text classification by a neural network*. Proceedings of the 23rd Annual Summer Computer Simulation Conference, Baltimore, USA, pp. 313-318, 1991.

[6] Jacobs P. S. *Joining statistics with NLP for text categorization*. Proceedings of ANLP-92, 3rd Conference on Applied Natural Language Processing, Association for Computational Linguistics, Morristown, US, editor Marcia Bates and Oliviero Stock, pp. 178-185, 1992.

[7] Geutner P., Bodenhausen U., and Waibel A. *Flexibility Through Incremental Learning: Neural Networks for Text Categorization*. Proceedings of WCNN-93, World Congress on Neural Networks, pp. 24-27, 1993.

[8] Riloff E., and Lehnert W. *Information extraction as a basis for high-precision text classification*. ACM Transactions on Information Systems, No.3, Vol. 12, pp. 296-333, 1994.

[9] Wiener E. D., Pedersen J. O. and Weigend A. S., *Neural network approach to topic spotting*. Proceedings of SDAIR-95, 4th Annual Symposium on Document Analysis and Information Retrieval, pp. 317-332, 1995.

[10] Joachims T. *A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization*. (Computer Science Technical Report CMU-CS-96-118). Carnegie Mellon University, 1995.

[11] Hanauer D. *Integration of phonetic and graphic features in poetic text categorization judgements*, Poetics, vol. 23, No.5, pp. 363-380, 1996.

[12] Merkl D. *Text classification with self-organizing maps: Some lessons learned*. Neurocomputing, vol. 21, No. 1/3, pp. 61-77, 1998.

[13] Hoon J. F., Frank S. L., Kowalczyk W. and Van der Ham F. *Neural Network Identification of Poets Using Letter Sequences*. Literary and Linguistic Computing, 14, pp. 311-338, 1999.

[14] Statamatos S., Fakotakis N., and Kokkiniakis G. *Automatic text categorization in terms of genre and author*. Computational Linguistics, vol. 26, No. 4, pp. 471-495, 2000.

[15] Aizawa A. *Linguistic Techniques to Improve the Performance of Automatic Text Categorization*. Proceedings of NLPERS-01, 6th Natural Language Processing Pacific Rim Symposium, pp. 307-314, 2001.

9 2 ~ 4 ~ » es e »	5 ~ ~ 3 ~ 3 e e	1 1 e e e e e	2 3 3 ~ ~ ~ ~
--	--------------------------------------	---------------------------------	---------------------------------