

استعمال معامل الإقتران المتعدد لتمييز القصائد العربية لعينة من

الشعراء

الباحث/ ندى أحمد جمعة
الجهاز المركزي للأحصاء
وزارة التخطيط

أ.م.د. مها عبد الكريم حمود
قسم هندسة الإنتاج والمعادن
الجامعة التكنولوجية

أ.د. عبد المنعم صالح رحمه
قسم علوم الحاسوب
الجامعة التكنولوجية

maha_alrawi@yahoo.com monem.rahma@yahoo.com
nadaalzubaidy@yahoo.com

تاريخ التقديم: 2018/3/28

تاريخ القبول: 2018/6/6

المستخلص

يهدف البحث إلى التمييز بين القصائد لكل شاعر من خلال خصائص وصفات الحروف العربية حيث إستعملت أربع تصنيفات للحروف العربية على أساسها تم إدراج تكرارات الحروف في جدول إقتران متعدد الإتجاهات (Multidimensional contingency table) ولكل إتجاه (تصنيف) مستويين أو أكثر ومنه أيجاد قيمة معامل الإقتران لكل قصيدة ضمن عينة البحث والتي تتألف من ستة شعراء من عصور مختلفة ولكل شاعر خمس قصائد.

تمت برمجة الطريقة باستخدام برنامج الماتلاب (MATLAB) وقد أثبتت الطريقة كفاءتها بنسبة 53% بالنسبة لعينة البحث المقترحة ككل وبنسبة تتراوح بين 90% إلى 95% بالنسبة للقصائد لكل شاعر.

المصطلحات الرئيسية للبحث/ الشعر العربي، أشكال الحروف العربية، صفات الحروف العربية، أل التعريف، جدول الإقتران المتعدد، مقاييس الإقتران.



مجلة العلوم

الاقتصادية والإدارية

العدد 108 المجلد 24

الصفحات 428.438



1. المقدمة (Introduction)

تبرز أهمية هذا البحث في إستخلاص صفات وخصائص الحروف العربية ضمن نص من القصائد العربية وتحليلها بطريقة تهدف إلى وضع مقاييس لنصوص القصائد المستعملة في عينة البحث للتمييز فيما بينها كما يمكن أن يمثل هذا البحث بذرة أساسية لإيصال اللغة العربية إلى مستوى أعلى وأقوى مما هي عليه في عالم الحوسبة وتكنولوجيا المعلومات بسبب ضعف البحوث الداعمة أو السائدة للغة العربية ولاسيما في مجال البحث (Searching) عند إستعمال متصفحات الأنترنت أو قد تكاد تكون شبه محدودة مقارنة باستخدام اللغة الانكليزية إذ تحتاج اللغة العربية فترة من الزمن لمواكبة بعض اللغات الأخرى ذات المحتوى العالي في مجال تكنولوجيا المعلومات.

يهدف البحث الى إقتراح طريقة لتمييز القصائد العربية اعتماداً على بعض خصائص الحروف العربية مثل (شكل الحرف من موقعه في الكلمة، قوة وضعف الحرف، الحروف الشمسية والقمرية، الحروف السامية والروادف) وذلك عن طريق اعتماد عينة من قصائد عدد من الشعراء يمثلون عصور أدبية مختلفة وإستعمالها في نظام يميز ويعطي قيم رقمية تطبق وفق مقاييس احصائية يمكن من خلال هذه القيم الناتجة إستحصاف فروقات وتشابهات تمكن من تمييز شاعر معين عن بقية الشعراء وايضا تعطي نسبة تقارب وتشابه شعر الشاعر نفسه.

في عام (2007) قام الباحث علاء حسين عيسى بتحديد طبقة القصيدة العربية من الشعر ذي الشطرين وعصرها الادبي بإستعمال نظام مقترح يعتمد على المقاييس الاحصائية ومعيار تكرارات الحروف للقصيدة.^[5]
في عام (2011) قدم الباحثان عبد المنعم صالح و لمى فايق خوارزمية مقترحة لرسم أو توليد صورة شعرية ثلاثية الأبعاد من النص الشعري العربي لتمييز بين القصائد.^[8]

في عام (2006) قدم الباحثون منصور بن محمد الغامدي ومحمد خورشيد ومصطفى الشافعي وآخرون طريقة للتشكيل الالي للنص العربي بأستعمال إنموذج ماركوف الخفي للتعرف على الكلام آليا.^[1]
في عام (2008) قدم الباحثان DARA AND TIZHOOSH SAHBA طريقة لتمييز وتصنيف النصوص الشعرية بأستعمال naïve Bayesian & Multiplayer Perceptron (MLP).^[11]

2. تصنيف الحروف العربية (Classification of Arabic letters)

تم في هذا البحث إستعمال عدة تقسيمات أو تصنيفات للحروف العربية لإستخدامها لاحقاً كصفات في عملية التمييز للنصوص الشعرية إذ تم إستعمال اربع تصنيفات للحروف وكل تصنيف بعدد من المستويات وهي كالاتي:

1-2. تقسيم الحروف على أساس موقعها من الكلمة

تأخذ الحروف العربية عدة أشكال إذ تم تقسيم الحروف بحسب موقعها من الكلمة من حيث وقوع الحرف في بداية الكلمة او في وسطها أو في آخرها (أي بحسب شكل الحرف) وعد هذا التقسيم الأساس في عملية التمييز للنصوص الشعرية لهذا البحث وتضمن هذا التصنيف أربع مجاميع هي^[3]:

- مجموعة الحروف الابتدائية وتشمل على الحروف بالأشكال (ب، ت، ث، ج، ح، خ، د، ذ، ر، ز، ط، ظ، ع، غ، ف، ق، ك، ل، م، ن، هـ، ي، نـ).
- مجموعة الحروف الوسطية وتشمل على الحروف بالأشكال (ب، ت، ث، ج، ح، خ، د، ذ، ر، ز، ط، ظ، ع، غ، ف، ق، ك، ل، م، ن، هـ، ي، نـ).
- مجموعة الحروف النهائية وتشمل على الحروف بالأشكال (أ، ب، ت، ث، ج، ح، خ، د، ذ، ر، ز، س، ش، ص، ض، ط، ظ، ع، غ، ف، ق، ك، ل، م، ن، هـ، ي، نـ، أ، إ، ي، آ).
- مجموعة الحروف المنفصلة وتشمل على الحروف بالأشكال (أ، ب، ت، ث، ج، ح، خ، د، ذ، ر، ز، س، ش، ص، ض، ط، ظ، ع، غ، ف، ق، ك، ل، م، ن، هـ، ي، نـ، أ، إ، ي، آ).



2-2. تقسيم الحروف على أساس القوة والضعف

يعتمد هذا التقسيم على صفات الحروف عند نطقها وتشمل على الصفات القوية (الجهر، الشدة، الإستعلاء، الإطباق، الإصمات، الصغير، القلقة، الإنحراف، التكرير، التفشي، الإستطالة) والصفات الضعيفة (الهمس، الرخاوة، الإستفال، الإنفتاح، الإذلاق، اللين، الخفاء) وتضمن هذا التصنيف خمسة مجاميع هي [2]:

- الحروف الأقوى: وهي الحروف التي كل صفاتها قوية أو بها صفة واحدة فقط ضعيفة (الضاد والطاء والظاء).

- الحروف القوية: وهي الحروف التي صفاتها القوية أكثر من صفاتها الضعيفة (الهمزة والجيم والدادال والغين والصاد والقاف).

- الحروف المتوسطة: وهي التي تتساوى فيها صفاتها القوية والضعيفة (الباء والزاي والراء والعين).
- الحروف الضعيفة: وهي صفاتها الضعيفة أكثر من صفاتها القوية (السين والكاف والتاء والألف واللام والشين والياء والحاء والذال والواو).

- الحروف الأضعف: وهي الحروف التي كل صفاتها ضعيفة أو بها صفة واحدة قوية (الفاء والحاء والثاء والهاء والميم والنون).

2-3. تقسيم الحروف على أساس (ال) التعريف

أل التعريف هي لام ساكنة زائدة عن بنية الكلمة مسبقة بهمزة وصل تكون اللام فيها على حالتين الأولى مظهرة والثانية مدغمة بحسب ما يقع بعدها من الحروف الهجائية وهي [2]:

- حالة الإظهار (الإظهار القمري) وسميت لام قمريه بسبب نطقها في كلمة (القمر) وتظهر مع الحروف (الهمزة أو الألف والياء والغين والحاء والجيم والكاف والواو والحاء والفاء والعين والقاف والياء والميم والهاء) وعددها أربعة عشر حرفاً مجموعة في القول (ابغ حجك وخف عقيمة).

- حالة الإدغام (الإظهار الشمسي) وسميت لام شمسية بسبب عدم نطقها في كلمة (الشمس) وتظهر مع الحروف (الطاء والثاء والصاد والراء والتاء والضاد والذال والنون والدادال والسين والظاء والزاي والشين واللام) وعددها أربعة عشر حرفاً وهي في أوائل البيت الشعري (طب ثم صل رحما تفض ضف ذا نعم دع سوء ظن زر شريفاً للكرم).

2-4. تقسيم الحروف على أساس الترتيب الأبجدي

وهو تقسيم منسوب إلى الساميين (أصحاب اللسان العبري) حيث تتشابه حروف اللغتين العبرية والعربية وتزيد العربية ببعض الحروف وهي [4]:

- الحروف السامية أو الأصول (الهمزة أو الألف والياء والجيم والدادال والهاء والواو والزاي والحاء والطاء والياء والكاف واللام والميم والنون والسين والعين والفاء والصاد والقاف والراء والشين والتاء) وهي مجموعة في المقطع (أبجد هوز حطي كلمن سعفص قرشت).

- الحروف الروادف (الثاء والحاء والذال والضاد والظاء والغين) وهي مجموعة في المقطع (تخذ ضطغ).

3. معامل الإقتران المتعدد (Multi contingency coefficient)

حين تكون بيانات الدراسة على شكل متغيرات نوعية يقسم كل متغير إلى مستويين أو أكثر يتم استخدام جداول الإقتران المتعدد لقياس الاستقلالية بين متغيرين أو أكثر باستخدام إحصاءة مربع كاي لبيرسون (Pearson chi-squared statistics) من خلال حساب لتكرارات التي يشرط أن تكون مستقلة [6]:

تم في هذا البحث استعمال جدول إقتران مكون من أربع إتجاهات (r x c x m x n) إذ إن (r) تمثل متجه الصفوف و(c) تمثل متجه الأعمدة و(m) تمثل الطبقة الثالثة و(n) الطبقة الرابعة، ويتم توزيع التكرارات المشاهدة على هذا الجدول ويرمز لها (f_{ijkl}) إذ إن (i=1,2, ..., r) و(j=1,2, ..., c) و(k=1,2, ..., m) و(l=1,2, ..., n) ومن ثم يتم حساب مجموع التكرارات لمستويات كل متجه كما في الصيغ الآتية [7]:



$$f_{i...} = \sum_{j=1}^c \sum_{k=1}^m \sum_{l=1}^n f_{ijkl} \quad \dots (1)$$

$$f_{.j..} = \sum_{i=1}^r \sum_{k=1}^m \sum_{l=1}^n f_{ijkl} \quad \dots (2)$$

$$f_{..k.} = \sum_{i=1}^r \sum_{j=1}^c \sum_{l=1}^n f_{ijkl} \quad \dots (3)$$

$$f_{...l} = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^m f_{ijkl} \quad \dots (4)$$

يتم تعريف القيم المتوقعة للتكرارات (Expected frequency) ويرمز لها (E_{ijkl}) من خلال فرضية العدم التي تنص على انه (ليس هناك اقتران معنوي بين المتغيرات أي ان المتغيرات مستقلة عن بعضها) والفرضية البديلة (هناك إقتران معنوي بين المتغيرات أي إن المتغيرات غير مستقلة عن بعضها) ويتم حسابها من الصيغ الآتية^[10]:

$$E_{ijkl} = N \cdot \frac{f_{i...}}{N} \cdot \frac{f_{.j..}}{N} \cdot \frac{f_{..k.}}{N} \cdot \frac{f_{...l}}{N} \quad \dots (5)$$

$$E_{ijkl} = \frac{f_{i...} \cdot f_{.j..} \cdot f_{..k.} \cdot f_{...l}}{N^3} \quad \dots (6)$$

ومن ثم يتم احتساب إحصاءة مربع كاي لبيرسون (Pearson chi-squared statistic) ويرمز لها (X^2) كما في الصيغة الآتية^[9]:

$$X^2 = \sum_i \sum_j \sum_k \sum_l \frac{(f_{ijkl} - E_{ijkl})^2}{E_{ijkl}} \quad \dots (7)$$

وبدرجة حرية (degree of freedom) تساوي $(df = r \cdot c \cdot k \cdot l - r - c - k - l + 3)$.

ومن خلال إحصاءة مربع كاي يتم احتساب معامل الإقتران المتعدد (C.C.) من الصيغة المقترحة من قبل العالم (Karl Pearson 1904) كالآتي^[7]:

$$C.C. = \sqrt{\frac{X^2}{X^2 + N}} \quad \dots (8)$$

إذ أن:

X^2 : قيمة مربع كاي المحسوبة.

N: حجم العينة.

تتراوح قيمة معامل الاقتران بين (1,0) فإذا كان قريب من الصفر دل ذلك على عدم وجود اقتران وإذا كان واحد دل ذلك على وجود علاقة بين المتغيرات المدروسة.



4. التطبيق العملي (Practical application)

تم تطبيق ما جاء في الجانب النظري على قصائد شعرية عربية لستة شعراء من عصور مختلفة ولكل شاعر خمسة قصائد وبحسب التصانيف للحروف العربية المذكورة فإن جدول الإقتران المتعدد سيتكون من أربع إتجاهات الأول يمثل الحروف بحسب موقعها من الكلمة بأربع مستويات والثاني الحروف بحسب القوة والضعف بخمس مستويات والثالث الحروف حسب ال التعريف بمستويين والرابع الحروف بحسب الترتيب الأبجدي بمستويين. كما تمت برمجة الطريقة المقترحة بالإستعانة ببرنامج الماتلاب (MATLAB R2014a) وبحسب الخطوات الآتية:

1. تم ترميز الحروف العربية للنص الشعري بحسب شكل الحرف (موقعه من الكلمة) من حيث كونه حرف ابتدائي أو حرف وسطي أو حرف نهائي أو حرف منفصل وكما موضح بالجدول (1).

جدول (1) ترميز أشكال الحروف العربية.

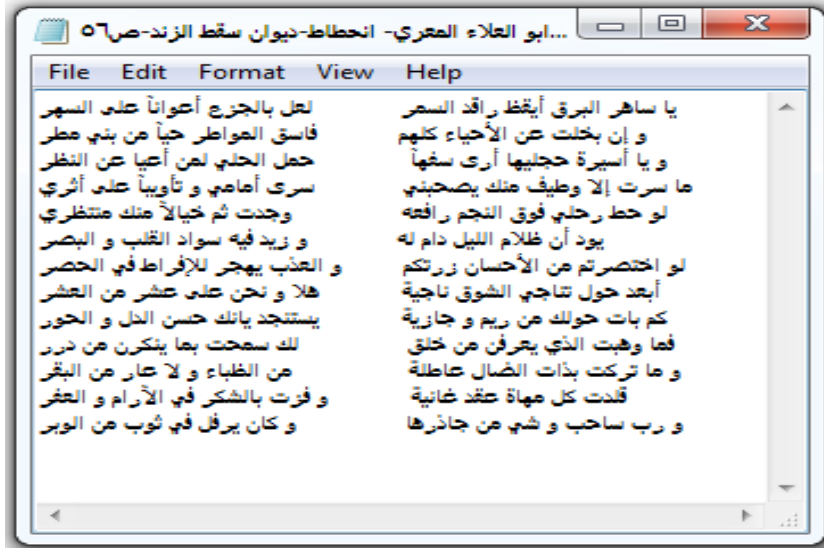
No	Letter X	First Letter	R1	Middle Letter	R2	Final Letter	R3	Separate d Letter	R4
1	ء	-	-	-	-	-	-	ء	D11
2	ا	-	-	-	-	ا	C12	ا	D12
3	ب	ب	A13	ب	B13	ب	C13	ب	D13
4	ت	ت	A14	ت	B14	ت	C14	ت	D14
5	ث	ث	A15	ث	B15	ث	C15	ث	D15
6	ج	ج	A16	ج	B16	ج	C16	ج	D16
7	ح	ح	A17	ح	B17	ح	C17	ح	D17
8	خ	خ	A18	خ	B18	خ	C18	خ	D18
9	د	-	-	-	-	د	C19	د	D19
10	ذ	-	-	-	-	ذ	C20	ذ	D20
11	ر	-	-	-	-	ر	C21	ر	D21
12	ز	-	-	-	-	ز	C22	ز	D22
13	س	س	A23	س	B23	س	C23	س	D23
14	ش	ش	A24	ش	B24	ش	C24	ش	D24
15	ص	ص	A25	ص	B25	ص	C25	ص	D25
16	ض	ض	A26	ض	B26	ض	C26	ض	D26
17	ط	ط	A27	ط	B27	ط	C27	ط	D27
18	ظ	ظ	A28	ظ	B28	ظ	C28	ظ	D28
19	ع	ع	A29	ع	B29	ع	C29	ع	D29
20	غ	غ	A30	غ	B30	غ	C30	غ	D30
21	ف	ف	A31	ف	B31	ف	C31	ف	D31
22	ق	ق	A32	ق	B32	ق	C32	ق	D32
23	ك	ك	A33	ك	B33	ك	C33	ك	D33
24	ل	ل	A34	ل	B34	ل	C34	ل	D34
25	م	م	A35	م	B35	م	C35	م	D35
26	ن	ن	A36	ن	B36	ن	C36	ن	D36
27	ه	ه	A37	ه	B37	ه	C37	ه	D37
28	و	-	-	-	-	و	C38	و	D38



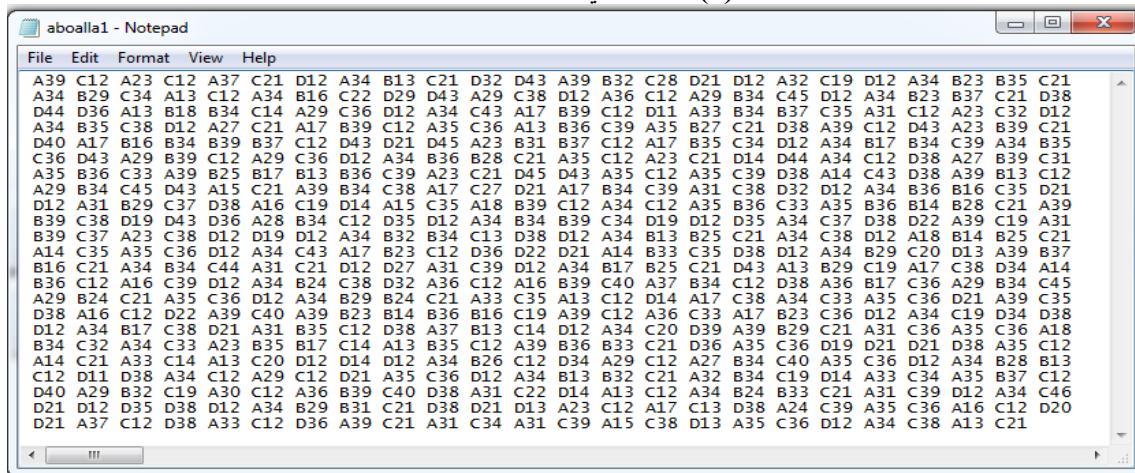
استعمال معامل الاقتران المتعدد لتمييز القصائد العربية لعينة من الشعراء

29	ي	ب	A39	ب	B39	سي	C39	ي	D39
30	ة	-	-	-	-	ة	C40	ة	D40
31	و	-	-	-	-	و	C41	و	D41
32	ئ	ن	A42	ن	B42	ئ	C42	ئ	D42
33	أ	-	-	-	-	أ	C43	أ	D43
34	إ	-	-	-	-	إ	C44	إ	D44
35	ى	-	-	-	-	ى	C45	ى	D45
36	آ	-	-	-	-	آ	C46	آ	D46

2. تم ترميز النص الشعري العربي ومثالاً على ذلك القصيدة الشعرية لأبو العلاء المعري الموضحة بالشكل (1) حسب جدول (1) وكما موضح بالشكل (2) الذي يبين النص المرمز. شكل (1) قصيدة شعرية لأبو العلاء المعري.



شكل (2) ملف نصي للحروف المرمزة.





استعمال معامل الإقتران المتعدد لتمييز القصائد العربية لعينة من الشعراء

3. حساب التكرارات للحروف بحسب التصنيفات الموضحة سابقا للحروف العربية ولكل مستوى.
4. توليد جدول الإقتران المتعدد (4*2*5*4) لكل قصيدة شعرية ومن ثم توزيع التكرارات المشاهدة للحروف على خلايا الجدول وكل حرف بحسب المستويات للتصانيف التي يقع ضمنها كما موضح بالشكل (3) والذي يبين هيكلية جدول الإقتران المتعدد لقصيدة أبو العلاء المعري.
- شكل (3) مثال على هيكلية جدول الإقتران المتعدد (التكرارات المشاهدة).

Multidimensional contingency table (4*5*2*2), poem 1

		1					2							
		1					2							
D4	D3	1					2							
	D2	D1	1	2	3	4	5	D2	D1	1	2	3	4	5
1	1	1	3	0	0	52	5	1	1	0	7	20	21	46
	2	2	1	4	0	29	10	2	2	0	9	14	15	21
	3	3	1	8	29	13	16	3	3	0	7	2	71	11
	4	4	1	4	16	11	5	4	4	0	16	4	58	3
	5	5						5	5					

		1					2							
		1					2							
D4	D3	1					2							
	D2	D1	1	2	3	4	5	D2	D1	1	2	3	4	5
2	1	1	1	0	0	0	3	1	1	0	1	0	3	0
	2	2	4	0	0	0	0	2	2	0	0	0	1	0
	3	3	1	0	0	3	0	3	3	0	0	0	0	0
	4	4	0	0	0	1	0	4	4	0	0	0	0	0
	5	5						5	5					

5. احتساب قيم التكرارات المتوقعة لجدول الإقتران بحسب الصيغة (6) ولكل قصيدة شعرية كما موضح بالشكل (4) الذي يبين القيم المتوقعة للتكرارات المشاهدة لقصيدة أبو العلاء المعري.
- شكل (4) جدول التكرارات المتوقعة.

Expected frequency table (4*5*2*2), poem 1

		1					2							
		1					2							
D4	D3	1					2							
	D2	D1	1	2	3	4	5	D2	D1	1	2	3	4	5
1	1	1	1.37	6.39	9.7	31.7	13.7	1	1	2.04	9.54	14.5	47.4	20.4
	2	2	0.91	4.26	6.46	21.1	9.13	2	2	1.36	6.36	9.65	31.6	13.6
	3	3	1.37	6.39	9.7	31.7	13.7	3	3	2.04	9.54	14.5	47.4	20.4
	4	4	1.01	4.69	7.12	23.3	10.1	4	4	1.5	7.01	10.6	34.8	15
	5	5						5	5					

		1					2							
		1					2							
D4	D3	1					2							
	D2	D1	1	2	3	4	5	D2	D1	1	2	3	4	5
2	1	1	0.05	0.22	0.33	1.07	0.46	1	1	0.07	0.32	0.49	1.6	0.69
	2	2	0.03	0.14	0.22	0.71	0.31	2	2	0.05	0.21	0.33	1.07	0.46
	3	3	0.05	0.22	0.33	1.07	0.46	3	3	0.07	0.32	0.49	1.6	0.69
	4	4	0.03	0.16	0.24	0.79	0.34	4	4	0.05	0.24	0.36	1.17	0.51
	5	5						5	5					



استعمال معامل الإقتران المتعدد لتمييز القصائد العربية لعينة من الشعراء

6. إحتساب قيمة مربع كاي حسب الصيغة (7) ولكل قصيدة حيث تمثل (N) عدد حروف القصيدة.
7. إحتساب قيمة معامل الإقتران حسب الصيغة (8) ولكل قصيدة والجدول (2) يوضح النتائج.
8. إحتساب المعدل لمعاملات الإقتران ذات القيم المتقاربة لقصائد كل شاعر.
9. إيجاد حدود الثقة الدنيا والعليا لكل معدلات معاملات الإقتران لتوضيح الفترة التي تقع ضمنها معاملات إقتران قصائد كل شاعر.

جدول (2) قيم معامل الإقتران

الحالة	الحد الأدنى $\alpha=0.05$	الحد الأعلى $\alpha=0.95$	معدل معامل الإقتران	معامل الإقتران	القصائد	الشاعر
نجاح	0.764	0.804	0.784	0.78	القصيدة 1	ابو العلاء المعري
نجاح				0.79	القصيدة 2	
نجاح				0.8	القصيدة 3	
فشل				0.77	القصيدة 4	
فشل				0.67	القصيدة 5	
فشل	0.701	0.743	0.722	0.72	القصيدة 1	علي بن أبي طالب (رض)
نجاح				0.7	القصيدة 2	
فشل				0.73	القصيدة 3	
فشل				0.73	القصيدة 4	
فشل				0.85	القصيدة 5	
نجاح	0.803	0.838	0.82	0.81	القصيدة 1	بدر شاكر السياب
نجاح				0.82	القصيدة 2	
نجاح				0.83	القصيدة 3	
فشل				0.72	القصيدة 4	
فشل				0.72	القصيدة 5	
فشل	0.715	0.747	0.731	0.72	القصيدة 1	حاتم الطائي
فشل				0.72	القصيدة 2	
فشل				0.73	القصيدة 3	
فشل				0.75	القصيدة 4	
فشل				0.74	القصيدة 5	
نجاح	0.68	0.694	0.687	0.69	القصيدة 1	كعب بن زهير
نجاح				0.68	القصيدة 2	
نجاح				0.69	القصيدة 3	
نجاح				0.69	القصيدة 4	
فشل				0.79	القصيدة 5	
نجاح	0.745	0.779	0.762	0.75	القصيدة 1	صفي الدين الحلي
نجاح				0.75	القصيدة 2	
نجاح				0.77	القصيدة 3	
نجاح				0.75	القصيدة 4	
نجاح				0.78	القصيدة 5	



ومن خلال إحتساب الكفاءة لكل شاعر (الكفاءة تساوي عدد القصائد ذات معامل إقتران متجانس مقسوم على العدد الكلي للقصائد مضروباً بالنسبة المئوية) ، يلاحظ من الجدول (2) إن قصائد الشاعر بدر شاكر السياب ثلاث منها تتراوح قيم معامل الإقتران لها على التوالي (0.83-0.81) وإثنين من القصائد معامل الإقتران (0.72) ما يدل على كفاءة في التمييز للقصائد بنسبة 60% تقريبا، وبالنسبة للشعراء علي بن أبي طالب (كرم الله وجهه) وأبو العلاء المعري وكعب بن زهير فهناك أربع قصائد لكل شاعر تعطي نفس قيمة معامل الإقتران تقريبا وتتراوح القيم على التوالي (0.73-0.70)، (0.80-0.77) (0.69-0.68) مما يعطي كفاءة في تميز القصائد بنسبة 80% تقريبا، أما بالنسبة للشعراء حاتم الطائي وصفي الدين الحلي فإن جميع قصائدهم أعطت نفس قيم معامل الإقتران تقريبا لكل شاعر والذي يتراوح على التوالي (0.75-0.72)، (0.75-0.72) (0.78) مما يعطي كفاءة في التمييز للقصائد بنسبة 100% تقريبا

أما بالنسبة لعينة القصائد الشعرية ككل تم أخذ معدل معامل الإقتران للقصائد التي تتقارب قيم معامل الإقتران لها ولكل شاعر كما موضح بالعمود الأخير من الجدول (2)، حيث تمت المقارنة على أساس التشابه بين قيم معامل الإقتران للقصائد و يلاحظ وجود تمييز للقصائد بين الشعراء إذ أعطت النتائج معدل معامل إقتران مختلف لكل شاعر و بفروقات واضحة نسبياً بينها ماعدا معدلات الشاعر علي بن أبي طالب (كرم الله وجهه) والشاعر حاتم الطائي حيث تتقارب قيم معدل معامل الإقتران مما يؤثر على كفاءة التمييز، وبشكل عام فإن التمييز واضح نسبياً بين عينة القصائد للشعراء.

يمكن إستنتاج إن الإختلافات في معاملات الإقتران للقصائد ولكل شاعر نابعة من إن عينة الشعراء للبحث من عصور وأزمان مختلفة وكذلك إختلاف في طبيعة ونوع قصائدهم فالشاعر بدر شاكر السياب من شعراء العصر الحديث وكتب في الشعر الحر لذا فإن مقاييس الإقتران له تختلف تماما عن بقية الشعراء مما تعطي تمييز جيد له ولقصائده، كذلك فإن الشاعر أبو العلاء المعري من العصر العباسي والقصائد من نوع المديح، والشاعر صفي الدين الحلي من عصر الإنحطاط حيث كان الشعر في هذا العصر يعكس السطحية والعامية، والشاعر كعب بن زهير من العصر الإسلامي وقصائده من النوع الغزلي جميعهم تم تمييزهم من خلال قيم مختلفة لمقاييس الإقتران.

أما الشاعر علي بن أبي طالب من العصر الإسلامي الذي أظهرت قصائده صفات ومزايا الإسلام ومدح الإله، والشاعر حاتم الطائي من العصر الجاهلي أظهرت قصائده صفات ومزايا عشيرته، فإن مقاييس الإقتران لهما قد أظهرت تقارب بينهما ويمكن أن يعزى السبب في تشابه المصطلحات والكلمات التي إستعملها في قصائدهم للعناية نفسها.

5. الإستنتاجات (Conclusion) :

تم من خلال تطبيق الطريقة المقترحة التوصل الى عدد من الأستنتاجات اهمها:

1. إن آلية إستعمال معامل الإقتران المتعدد غير شائعة كثيراً لا سيما إستعمالها في تمييز النصوص العربية التي تساعد في توسيع وتطوير وإدرج اللغة العربية في مجالات عديدة ولاسيما في المجالات التي تفتقر إلى إستعمال اللغة العربية كالإنترنت وبعض اللغات البرمجية.
2. أعطت الطريقة نسبة نجاح قدرها (53%) من عينة البحث وهي نسبة مقبولة نظراً لقلّة التعامل مع المحتوى العربي للنصوص المكتوبة باللغة العربية مقارنة باللغات الأخرى وقلّة الطرائق والبرامجيات التي تدعم اللغة العربية في مجال البحث.
3. ضمن عينة البحث فشل النظام في تحديد هوية الشاعر للقصيدة لبعض قصائد العينة بنسبة (47%) لان قيم معامل الإقتران لها تشابهت أو إقتربت من بعضها ويعزى ذلك الى تشابه نوع الشعر المستخدم في هذه القصائد ومن ثم تشابه الكلمات والمصطلحات.



المصادر (References)

1. الغامدي ، منصور ، خورشيد ، محمد ، الشافعي ، مصطفى ، " نظام حاسوبي لتشكيل النص العربي " ، 2006 ، صفحة 2.
2. المرصفي، عبد الفتاح السيد عجمي، 1980، "هداية القاري إلى تجويد كلام الباري"، الطبعة الثانية، مكتبة طيبة، المدينة المنورة ص. ب: 218 ، صفحة 93 و 204.
3. زرقة، أحمد، 1993، "أسرار الحروف"، الطبعة الأولى، دار الحصاد للنشر والتوزيع، دمشق ص. ب: 4490 ، صفحة 113.
4. صالح، فخري محمد، 1994، "اللغة العربية أداء ونطقا وإملاء وكتابة"، الطبعة الثانية، دار الوفاء، المنصورة، ص. ب: 230 ، صفحة 26.
5. عيسى ، علاء حسين ، 2007 ، " تحديد طبقة القصيدة العربية من الشعر العربي ذي الشطرين و عمرها الأدبي عن طريق اعتماد الإحصاء كأحد تقنيات التنقيب عن البيانات " ، الجامعة التكنولوجية ، كلية الرشيد للهندسة والعلوم ، قسم علوم الحاسبات ، صفحة 13.
6. Chowdhury, Sadeq R. and Mcgilchrist, C. A., 2001, "Analysis of contingency tables with clustered observation", Aust. N. Z. J. Stat. 43(3), 351-358 ,page 3.
7. Everitt, B. S., 1977, "The analysis of contingency tables", Halsted press, a division of john Wiley & sons, inc., newyork, pages 64& 68.
8. Rahma , Abd Al_Monem , Faik , Luma , 2011 , "Implementation Of 3D Virtual Poetry Image By Using Bezier Cubic Splines And Open GL Primitives" , Eng.&Tech.Journal,Vol.29, No.12 , page 2364.
9. Shaffer, Juliet Oopper, 1973, "Defining and testing hypotheses in multidimensional contingency tables", Psychological Bulletin, vol. 79, no. 2, 127-141 , page 6.
10. Stemmler, Mark, 2014, "Person-centered methods", Springer Cham Heidelberg newyork Dordrecht London , page 5.
11. TIZHOOSH, SAHBA, AND DARA , 2008 , " Poetic Features for Poem Recognition: A Comparative Study " , JOURNAL OF PATTERN RECOGNITION RESEARCH (2008) 24-39 , page 27.



Using Multidimensional Contingency Coefficient to Discrimination Arabic Poems for Sample of Poets

Abstract

The purpose of this paper to discriminate between the poetic poems of each poet depending on the characteristics and attribute of the Arabic letters. Four categories used for the Arabic letters, letters frequency have been included in a multidimensional contingency table and each dimension has two or more levels, then contingency coefficient calculated.

The paper sample consists of six poets from different historical ages, and each poet has five poems. The method was programmed using the MATLAB program, the efficiency of the proposed method is 53% for the whole sample, and between 90% and 95% for each poet's poems.

Keywords: Arabic poetry, Arabic letters shapes, Arabic letters characteristics, Al-alta'reef, multidimensional contingency table, contingency measurements.