

פיתוח מדדים לאיתור חיבורים משורבטים (חיבורי ג'יבריש)



ענת בן-סימון ויעל שפרן
המרכז הארצי לבחינות ולהערכה

פברואר, 2015

הכנס ה-11 של האגודה הישראלית לפסיכומטריקה (אפי), ירושלים

הקדמה

- לאור השימוש ההולך וגובר בהערכה ממוחשבת של חיבורים ותרומתו לייעול הליך ההערכה, יש חשיבות רבה לשיפור הכלים המשמשים בהערכה זו.
 - היכולת לזהות חיבורים חריגים ולעמוד על המאפיינים התורמים לחריגותם צפויה לתרום רבות ל –
 - שיפור דיוק ההערכה
 - לתוקף הנראה של הערכה (באמצעות מתן משוב הולם לכותבים)
 - ליכולת לזהות חיבורי ג'יבריש יש תרומה דיפרנציאלית בהתאמה להקשר בו נכתב החיבור
-
- המוטיבציה למחקר הנוכחי נבעה מהצורך להעריך באופן אוטומטי חיבורים שנכתבו במסגרת מערכת מתוקשבת להתנסות עצמית בבחינה הפסיכומטרית ('פסיכורשת')

מטרת המחקר

פיתוח מדדים תקפים לאיתור חיבורים לא תקינים;

- חיבורים שנכתבו כלאחר יד ואינם עומדים בקריטריונים של כתיבה תקנית
- חיבורים שלא נכתבו לנושא הכתיבה

דוגמאות של חיבורים לא תקינים

חיבור -

- ❑ "ריק" או קצר מאוד
- ❑ הכולל רצפים מקריים של אותיות או מילים
- ❑ הכולל משפטים או פסקאות חוזרים
- ❑ שחלקו או רובו אינו כתוב בעברית
- ❑ שחלקו או כולו נכתב לא לנושא הכתיבה

שיטה:



קריטריונים לזיהוי חיבורי ג'יבריש

- הוגדרו 10 מדדים בעברית (8 בערבית) לזיהוי חיבורי ג'יבריש
- כל המדדים מבוססים על מאפייני טקסט כמותיים המופקים באופן ממוחשב (על ידי NiteRater)
- לכל מדד הוגדר ערך סף, לפיכך, הציון הממוחשב לכל מדד הוא דיכוטומי (1/0)
- בנוסף מחושב ציון כולל שהוא סכום ציוני המדדים

כלל החלטה:

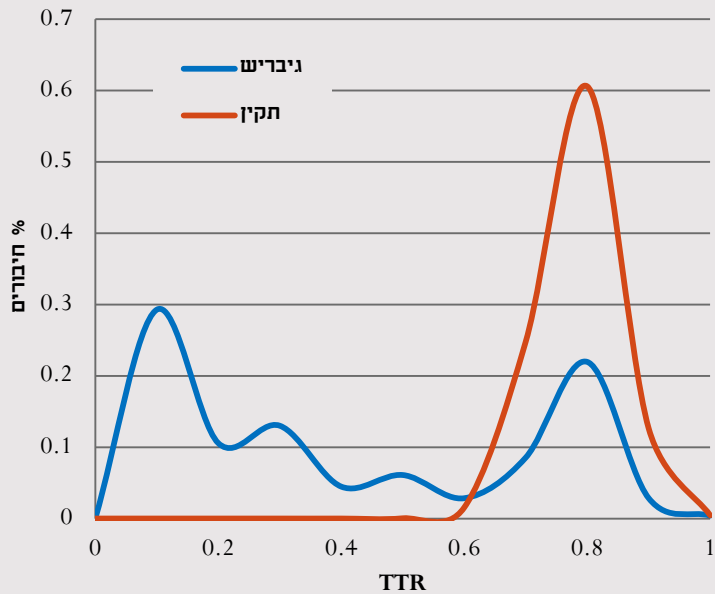
- חיבור ייחשב כג'יבריש אם החיבור חורג במדד אחד או יותר מערך הסף

רשימת הקריטריונים

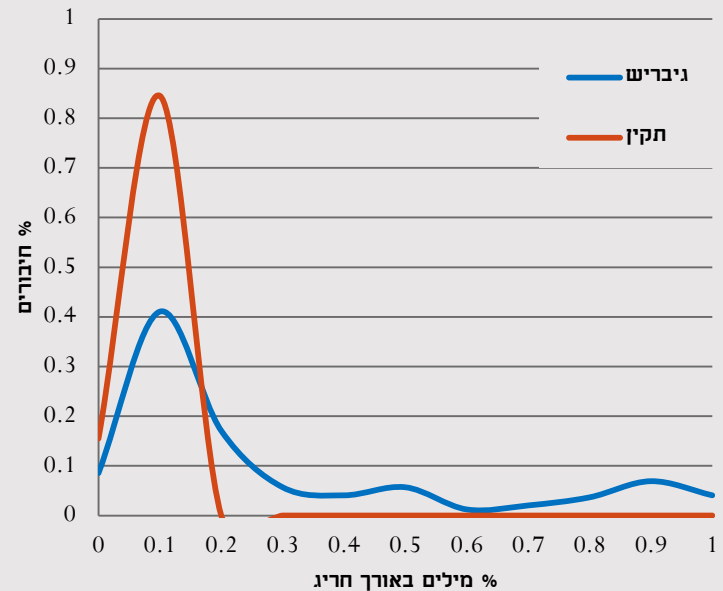
קריטריון (מדד)	תיאור	ערך סף
1. יחס תבנית-תמניות	% נמוך של type/token ratio (TTR)	<.30
2. משפטים חוזרים	משפטים שונים החוזרים על עצמם	3 משפטים לפחות
3. פסקאות קצרות	% גבוה של פסקאות קצרות מאד פחות מ-5 מילים	>.50
4. טקסט חופף גריין (העתקת הגריין)	% גבוה של מחרוזות / משפטים החופפים לגריין	>.80
5. אורך מילה	% גבוה של מילים באורך אחת או מעל 10 מילים	>.30
6. מילים זרות	% גבוה של מילים שאינן בעברית	>.50
7. רצפים לא חוקיים	רצפים של מילים לא חוקיות (למשל, מעל 10 תווים, מילים שלא קיימות במילון)	3 רצפים של 4 מילים
8. טקסט שאינו חופף גריין (טקסט לא לנושא)	% נמוך של מחרוזות / לקסמות חופפות גריין	<.05 <.015
9. מילים לא חוקיות בעברית	% מילים לא חוקיות (למשל, אות סופית בתחילת או באמצע מילה, אות תחילית בסוף מילה)	>.05
10. מילים לא במילון	% גבוה של מילים שאינן קיימות במילון	>.30

דוגמה: התפלגות חיבורי ג'יבריש וחיבורים תקינים בשני קריטריונים

יחס תבניות תמוניות (TTR)



אחוז מילים באורך 1 או 10 ויותר



שיטה: מדגם

5,196 חיבורים שנכתבו במסגרת 'פסיכורשת' □

מתוכם -

4,108 חיבורים בשפה העברית □

1,088 חיבורים בשפה הערבית □

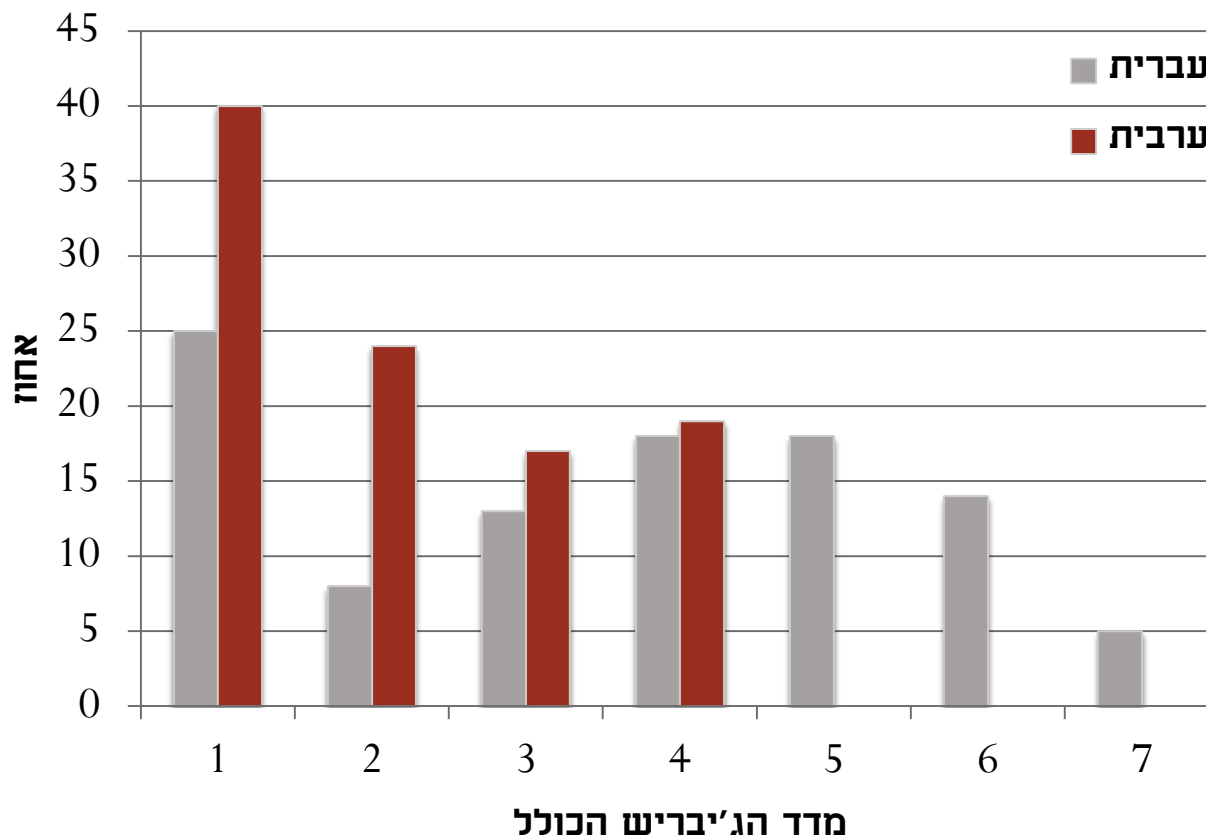
ניתוח הנתונים

- כל חיבור סווג באופן ממוחשב כתקין או חריג
- נחקרה ההתפלגות של חיבורי הג'יבריש והחיבורים התקינים בכל קריטריון ובציון הג'יבריש הכולל
- נבדק דיוק הסיווג -
 - עבור חיבורים שאורכם עלה על 100 מילים
 - סיווג של מעריך אנושי שימש כקריטריון תיקוף

התפלגות החיבורים לפי קטגוריות התקינות

סך הכול	קצרים	ג'יבריש	תקינים	
				עברית
4,108	1,143	246	2,719	N
	28.4	6.1	66.2	%
				ערבית
1,088	373	58	657	N
	34.3	5.3	60.4	%

התפלגות חיבורי ג'יבריש לפי הציון במדד הג'יבריש הכולל



שיעור חיבורי הג'יבריש ש"כשלו" בכל קריטריון

קריטריון	עברית	ערבית
.1 TTR	53	43
.2 משפטים חוזרים	13	10
.3 פסקאות קצרות	7	2
.4 טקסט חופף גריין במרביתו	11	9
.5 אורך מילה	28	31
.6 מילים זרות	0	2
.7 רצפים לא חוקיים	60	52
.8 טקסט שאינו חופף גריין	75	67
.9 מילים לא חוקיות בשפה	44	-
.10 מילים לא במילון	34	-

התפלגות החיבורים לפי דיוק הזיהוי

חיבורים בשפה העברית (N=2,965)

\leq שיעור הדיוק = 99.6%

סיווג ממוחשב			
ג'יבריש	תקין		סיווג אנוש
0 (false alarm)	2,708 (correct rejection)	תקין	
246 (hit)	11 (miss)	ג'יבריש	

חיבורים בשפה הערבית (N=270)

\leq שיעור הדיוק = 96.7%

סיווג ממוחשב			
ג'יבריש	תקין		סיווג אנוש
1	204	תקין	
57	8	ג'יבריש	

איתור חיבורים לא לנושא

עברית

- נדגמו 100 טקסטים אקראיים ממקורות שונים (עיתונות, ויקיפדיה וחיבורים שנכתבו לנושא שונה מנושא החיבור בפסיכורשת)
- 78% מהחיבורים אותרו בהצלחה

ערבית

- נדגמו 1,000 חיבורים שנכתבו במסגרת הבחינה הפסיכומטרית לנושא שונה מנושא החיבור בפסיכורשת
- 99% מהחיבורים אותרו בהצלחה

<.05
<.015

% נמוך של מחרוזות / לקסמות חופפות גריין

8. טקסט שאינו חופף גריין
(טקסט לא לנושא)

נושאי החיבור בפסיכורשת

בעבר במרבית מדינות העולם ניתנה הזכות להצביע בבחירות לפרלמנט בגיל 21, אך זה כמה עשורים ניכרת מגמה עולמית להוריד את גיל הבחירה ל-18. בכמה מדינות, ובהן ישראל, הוגשו הצעות חוק להורדת גיל הבחירות מ-18 ל-17 ואף ל-16, ויש מדינות שכבר עשו כן. בדיון הציבורי על הורדת גיל הבחירה נשקלות שאלות כגון: מאיזה גיל אדם זכאי להשפיע על הרכב הפרלמנט, המקבל החלטות שמשפיעות השפעה ישירה על חייו? האם בני הנוער בשלים דיים להחליט החלטות המשפיעות על הציבור כולו? מה יהיו ההשלכות של הכללת בני הנוער בקבוצת הבוחרים על המעורבות של בני הנוער בנעשה במדינה ועל המערכת הפוליטית בכללה?

**לדעתכם, האם יש להוריד את גיל הבחירה לכנסת? נמקו.
תוכלו להיעזר בשאלות המוצגות בפסקה.**

في دروس التاريخ في المدارس يتعلم الطلاب عن أحداث ومجريات حصلت في الماضي، عن الدوافع التي أدت إليها وعن تأثيراتها وتبعاتها. في معظم دول العالم موضوع التاريخ هو أحد المواضيع المركزية التي تُدرّس في المدارس.

לماذا من المهمّ تعلمّ التاريخ: كيف في رأيك يساهم تدريس التاريخ في نفع الإنسان والمجتمع؟

בשיעורי ההיסטוריה בבתי הספר לומדים על אירועים ועל תהליכים שהתרחשו בעבר, על הגורמים להם ועל השלכותיהם. ברוב מדינות העולם מקצוע זה הוא אחד המקצועות המרכזיים הנלמדים בבית הספר.

מדוע חשוב ללמד היסטוריה? כיצד לדעתכם לימודי ההיסטוריה תורמים לאדם ולחברה?

סיכום

□ האלגוריתם לזיהוי חיבורי ג'יבריש מדויק דיו וניתן להשתמש בו בגרסה התפעולית של פסיכורשת.

אבל -

- יש לבצע תיקוף חוזר של דיוק הזיהוי
- יש לבדוק את מידת הכללה של האלגוריתם
 - לנושאי כתיבה אחרים
 - לסוגות וסוגים שונים של כתיבה
- אין ספק שניתן להשביח את המדדים

תודה רבה!



The word gibberish comes from a Sufi mystic, whose name was Jabbar, and that was his only meditation.