**דוח מסקנות: בחינת שיפור הסיפורים ע"י המלצות המבקר האוטומטי**

**מוטיבציה לניסוי:**

כחלק ממטרות הפרויקט, התבצע ניסוי לבחינת השיפור בסיפורים המיוצרים לאחר יישום ההמלצות שהתקבלו ע"י המבקר האוטומטי שיצרנו במסגרת הפרויקט.

**אופן ביצוע הניסוי:**

* יצרנו 50 שלדי סיפור תקינים באמצעות כל אלגוריתם Plotto ו- Plot Genie .
* עבור כל שלד סיפור ביצענו הרחבה באמצעות Prompt שנשלח למודל Claude 3.7 Sonnet, על מנת שיפיק עבורנו סיפורים באורך של כ2000 מילים.
* כל אחד מבין 50 הסיפורים דורגו באמצעות המבקר האוטומטי בשיטה זהה.
* התקבל דירוג בצירוף הסברים והמלצות לשיפור הסיפור עבור כל אחד מהסיפורים, מהדירוגים יצרנו Prompt ייעודי שמטרתו לשפר כל הסיפור על סמך הביקורת.
* ביצענו את השיפור לכל אחד מהסיפורים באמצעות שליחתו למודל Claude 3.7 Sonnet.
* כל אחד מבין 50 הסיפורים המשופרים דורגו באמצעות המבקר האוטומטי פעם נוספת בשיטה זהה.
* כנספח לניסוי זה מצורפת טבלה ובה ציון כל סיפור עבור אחד מהמדדים והציון הכולל שקיבל הסיפור + הערכים הסטטיסטיים. ציוני המדדים בטבלה הם לאחר התחשבות במשקלן מכלל הציון הסופי, כלומר אם ציון מדד הדמויות בסיפור הוא 1.155 הציון שקיבל בפועל מתוך 4 נקודות הוא 1.155/0.3=3.85 כיוון שמשקל הקטגוריה הוא 30% מהציון הכולל.

**מהלך ההשוואה:**

בוצעה השוואה בין הגרסאות של כל אחד מהסיפורים (לפני ואחרי השיפור), כאשר המדדים להשוואה הם: דמויות, קונפליקט, מלאכת הכתיבה והיגיון.

עבור כל אחד מהמדדים השתמשנו בשאלות שהופיעו במאמר להערכת תסריטים תוך התאמת השאלות לסיפורים, [קישור למאמר](https://www.specscout.com/score).

**תוצאות ההשוואה:**

הושלמה השוואה בין 50 סיפורים עבור כל אלגוריתם בגרסתם המקורית לעומת גרסתם המשופרת לאחר יישום המלצות המבקר האוטומטי (איטרציה אחת של שיפור בלבד בשלב זה).

**טבלת ההשוואה עבור אלגוריתם Plotto:**

מתוך 50 הסיפורים, **ב־33 מהם** **נרשם שיפור** בציון הכולל. **ב־3 סיפורים לא חל שיפור**, וב־**14 סיפורים** אף נרשמה **ירידה קלה**.  
- ממוצע הציונים לפני השיפור: 3.643   
- ממוצע הציונים לאחר השיפור: 3.73   
- ממוצע השיפור הכולל: **0.087 נקודות**.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר | שם הסיפור | ציון לפני | ציון אחרי | שינוי בציון |
| 1 | Adrift in Betrayal | 3.7325 | 3.53 | -0.20 (-5.4%) |
| 2 | Between Two Worlds | 3.7475 | 3.5675 | -0.18 (-4.8%) |
| 3 | Breaking Free | 3.365 | 3.985 | 0.62 (18.4%) |
| 4 | Etchoes of Fidelity | 3.6275 | 3.56 | -0.07 (-1.9%) |
| 5 | Shadows of Loyalty | 3.53 | 3.605 | 0.07 (2.1%) |
| 6 | Shadows of Resemblance | 3.5975 | 3.5675 | -0.03 (-0.8%) |
| 7 | The Billiard General's Gambit | 3.7475 | 3.7775 | 0.03 (0.8%) |
| 8 | The Birthmark's Shadow | 3.7025 | 3.65 | -0.05 (-1.4%) |
| 9 | The Borrowed Muse | 3.7325 | 3.7775 | 0.04 (1.2%) |
| 10 | The Collector's Folly | 3.7675 | 3.7675 | 0.00 (0.0%) |
| 11 | The Courage Talisman | 3.71 | 3.71 | 0.00 (0.0%) |
| 12 | The Crimson Brotherhood | 3.4325 | 3.65 | 0.22 (6.3%) |
| 13 | The Detective's Gambit | 3.545 | 3.77 | 0.23 (6.3%) |
| 14 | The Diamond's Redemption | 3.65 | 3.8 | 0.15 (4.1%) |
| 15 | The Double Life | 3.6875 | 3.6725 | -0.02 (-0.4%) |
| 16 | The Gem's Gambit | 3.68 | 3.85 | 0.17 (4.6%) |
| 17 | The Ghost of Vengeance | 3.38 | 3.6875 | 0.31 (9.1%) |
| 18 | The Honest Impersonator | 3.59 | 3.5725 | -0.02 (-0.5%) |
| 19 | The Honorable Resistance | 3.74 | 3.62 | -0.12 (-3.2%) |
| 20 | The House of Cards | 3.7175 | 3.785 | 0.07 (1.8%) |
| 21 | The Illusion of Enterprise | 3.71 | 3.925 | 0.21 (5.8%) |
| 22 | The Keeper of Secrets | 3.6275 | 3.6575 | 0.03 (0.8%) |
| 23 | The Masquerade of Secrets | 3.7375 | 3.6875 | -0.05 (-1.3%) |
| 24 | The Method Actor's Delusion | 3.6275 | 3.68 | 0.05 (1.4%) |
| 25 | The Method Actor's Dilemma | 3.71 | 3.88 | 0.17 (4.6%) |
| 26 | The Miniature's Secret | 3.2 | 3.79 | 0.59 (18.4%) |
| 27 | The Monument of Truth | 3.6825 | 3.7325 | 0.05 (1.4%) |
| 28 | The Mysterious Artifact | 3.4175 | 3.7175 | 0.30 (8.8%) |
| 29 | The Oath Between Us | 3.8975 | 4 | 0.10 (2.6%) |
| 30 | The Phantom of Ideals | 3.7175 | 3.7175 | 0.00 (0.0%) |
| 31 | The Price of Ambition | 3.7175 | 3.6875 | -0.03 (-0.8%) |
| 32 | The Price of Deception | 3.6875 | 3.77 | 0.08 (2.2%) |
| 33 | The Price of Devotion | 3.6275 | 3.73 | 0.10 (2.8%) |
| 34 | The Price of Dreams | 3.9563 | 4 | 0.04 (1.1%) |
| 35 | The Price of Love | 3.755 | 3.85 | 0.09 (2.5%) |
| 36 | The Price of Loyalty | 3.6725 | 3.85 | 0.18 (4.8%) |
| 37 | The Shattered Mirror of Self-Deception | 3.47 | 3.4325 | -0.04 (-1.1%) |
| 38 | The Talisman's Truth | 3.635 | 3.7825 | 0.15 (4.1%) |
| 39 | The Ventriloquist's Transformation | 3.3425 | 3.425 | 0.08 (2.5%) |
| 40 | The Weight of Conscience | 3.7175 | 3.9775 | 0.26 (7.0%) |
| 41 | The Weight of Honor | 3.545 | 3.77 | 0.23 (6.3%) |
| 42 | The Price of Pride | 3.5525 | 4 | 0.45 (12.6%) |
| 43 | The Prodigal's Return | 3.7775 | 4 | 0.22 (5.9%) |
| 44 | The Reflection of Beauty | 3.68 | 3.71 | 0.03 (0.8%) |
| 45 | The Reflection of Desire | 3.71 | 3.74 | 0.03 (0.8%) |
| 46 | The Reluctant Impersonator | 3.7975 | 3.7475 | -0.05 (-1.3%) |
| 47 | The Secret Guardian | 3.485 | 3.5375 | 0.05 (1.5%) |
| 48 | Crossroads of Duty | 3.725 | 3.5675 | -0.16 (-4.2%) |
| 49 | Echoes of Fidelity | 3.6275 | 3.56 | -0.07 (-1.9%) |
| 50 | Rising from the Ashes | 3.6575 | 3.71 | 0.05 (1.4%) |

|  |  |
| --- | --- |
| **Median** | **0.15** |
| **STD** | **0.1511** |
| **Num Of Stories** | **50** |

**השיפורים בכל אחד מתחומי הביקורת:**

* **דמויות:** שיפור ממוצע של 3% – בולט בעיקר בהעמקת מאפיינים פנימיים, קונפליקטים רגשיים ותרומת דמויות המשנה לעלילה.
* **קונפליקט:** שיפור ממוצע של 1.11% – שיפור בעקביות הקונפליקט ובקשר הישיר לדמות הראשית ולמניעיה.
* **היגיון:** שיפור של 5.53% – סידור העלילה, סגירת קצוות פתוחים והפחתת סתירות פנימיות.
* **מלאכת הכתיבה:** שיפור מתון של 2.52% – שיפור בזרימה, בניסוח ובשימוש בדימויים תיאוריים.

**טבלת ההשוואה עבור אלגוריתם Plot Genie:**

מתוך 50 הסיפורים, **ב־35 מהם** **נרשם שיפור** בציון הכולל. **ב־2 סיפורים לא חל שיפור**, וב־**13 סיפורים** אף נרשמה **ירידה קלה**.  
- ממוצע הציונים לפני השיפור: 3.63   
- ממוצע הציונים לאחר השיפור: 3.702  
- ממוצע השיפור הכולל: **0.072 נקודות**.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר | שם הסיפור | ציון לפני | ציון אחרי | שינוי בציון |
| 1 | The Jade Collector's Dilemma | 3.8 | 3.8 | 0.00 (0.0%) |
| 2 | Stars in the Machine | 3.5525 | 3.7175 | 0.17 (4.6%) |
| 3 | Echoes in the Canyon | 3.6875 | 3.755 | 0.07 (1.8%) |
| 4 | Shadows of the Reservation | 3.6425 | 3.77 | 0.13 (3.5%) |
| 5 | Chains of Reform | 3.74 | 3.785 | 0.04 (1.2%) |
| 6 | The Navigator's Compass | 3.56 | 3.6725 | 0.11 (3.2%) |
| 7 | Whispers in the Pines | 3.6575 | 3.7175 | 0.06 (1.6%) |
| 8 | Depths of Loyalty | 3.6125 | 3.6575 | 0.04 (1.2%) |
| 9 | Valuation of the Heart | 3.635 | 3.6725 | 0.04 (1.0%) |
| 10 | The Philanthropist's Sacrifice | 3.215 | 3.5675 | 0.35 (11.0%) |
| 11 | Sands of Redemption | 3.6275 | 3.865 | 0.24 (6.5%) |
| 12 | Shadows of the Highlands | 3.7225 | 3.605 | -0.12 (-3.2%) |
| 13 | Foundations of Vengeance | 3.665 | 3.7625 | 0.10 (2.7%) |
| 14 | Shadows of the Magnolia Plantation | 3.5225 | 3.4625 | -0.06 (-1.7%) |
| 15 | Redemption at Pinewood Auto Camp | 3.6425 | 3.785 | 0.14 (3.9%) |
| 16 | The Uprising at Blackrock Mine | 3.5225 | 3.9325 | 0.41 (11.6%) |
| 17 | Shadows at the Lighthouse | 3.6425 | 3.625 | -0.02 (-0.5%) |
| 18 | Whispers in the Heather | 3.59 | 3.6275 | 0.04 (1.0%) |
| 19 | Echoes of the Canyon | 3.6575 | 3.5525 | -0.11 (-2.9%) |
| 20 | The Facets of Deception | 3.7175 | 3.7025 | -0.02 (-0.4%) |
| 21 | Caged Ambitions | 3.6125 | 3.74 | 0.13 (3.5%) |
| 22 | The Oarsman's Redemption | 3.6875 | 3.7175 | 0.03 (0.8%) |
| 23 | The Midnight Express | 3.7675 | 3.7375 | -0.03 (-0.8%) |
| 24 | The Vineyard's Gambit | 3.775 | 3.7475 | -0.03 (-0.7%) |
| 25 | The Arsenal's Perilous Alliance | 3.575 | 3.5675 | -0.01 (-0.2%) |
| 26 | The Dance of Duty and Redemption | 3.815 | 4 | 0.19 (4.8%) |
| 27 | Echoes of the Aurora | 3.575 | 3.515 | -0.06 (-1.7%) |
| 28 | Shadows of the Immortal Jungle | 3.4325 | 3.555 | 0.12 (3.6%) |
| 29 | Beneath the Cypress Shadows | 3.685 | 3.835 | 0.15 (4.1%) |
| 30 | Threads of Deception | 3.55 | 3.65 | 0.10 (2.8%) |
| 31 | Echoes of Vengeance | 3.6725 | 3.74 | 0.07 (1.8%) |
| 32 | Ripples of Vengeance | 3.7325 | 3.835 | 0.10 (2.7%) |
| 33 | Time's Gentle Remedy | 3.6875 | 3.5225 | -0.17 (-4.5%) |
| 34 | The House Always Wins | 3.6125 | 3.5375 | -0.08 (-2.1%) |
| 35 | The Timekeeper's Deception | 3.5075 | 3.9025 | 0.40 (11.3%) |
| 36 | Ashes of Vengeance | 3.635 | 3.8125 | 0.18 (4.9%) |
| 37 | Shadows of Redemption | 3.65 | 3.7175 | 0.07 (1.8%) |
| 38 | Whispers in the Reeds | 3.605 | 3.77 | 0.17 (4.6%) |
| 39 | Ripples of Redemption | 3.5975 | 3.6875 | 0.09 (2.5%) |
| 40 | The Minstrel's Redemption | 3.7475 | 3.755 | 0.01 (0.2%) |
| 41 | The Lancer's Oath | 3.775 | 3.6725 | -0.10 (-2.7%) |
| 42 | Currents of Courage | 3.755 | 3.7825 | 0.03 (0.7%) |
| 43 | The Masquerade of Loyalty | 3.62 | 3.7175 | 0.10 (2.7%) |
| 44 | Masks of Vengeance | 3.7675 | 3.7625 | 0.00 (-0.1%) |
| 45 | Harbor of Shadows | 3.365 | 3.23 | -0.14 (-4.0%) |
| 46 | Bitter Tides | 3.6275 | 3.7775 | 0.15 (4.1%) |
| 47 | The Peddler's Daughter | 3.6875 | 3.7675 | 0.08 (2.2%) |
| 48 | Ashes of Redemption | 3.6575 | 3.7475 | 0.09 (2.5%) |
| 49 | Flight of Fate | 3.38 | 3.6275 | 0.25 (7.3%) |
| 50 | The Mortician's Dilemma | 3.695 | 3.6425 | 0.00 (0.0%) |

|  |  |
| --- | --- |
| **Median** | **0.07** |
| **STD** | **0.1251** |
| **Num Of Stories** | **50** |

**השיפורים בכל אחד מתחומי הביקורת:**

* **דמויות:** שיפור ממוצע של 2.68% – בולט בהעשרת עומק רגשי, מיקוד במניעים הפנימיים והעמקת תרומת דמויות המשנה.
* **קונפליקט:** שיפור ממוצע של 3.39% – בניית קונפליקט רלוונטי, עקבי ומשמעותי לדמות הראשית.
* **היגיון:** שיפור של 5.33% – העלילה מאורגנת וברורה יותר, עם פחות סתירות ויותר קוהרנטיות.
* **מלאכת הכתיבה:** שיפור ממוצע של 2.49% – סגנון כתיבה מדויק יותר, שימוש עשיר בתיאורים ושיפור בקצב הקריאה.

**דוגמה לשיפור חלק "הדמויות" בעקבות ההמלצות האוטומטיות:**

**שם הסיפור : *Stars in the Machine***

**בסיפור זה נרשם שיפור בחלק הדמויות מציון של 1.035 ל1.155 . להלן כמה**

**דוגמאות שיתארו את השיפור בחלק זה.**

**בגרסה המקורית**, הקשרים החברתיים של אליאס, הדמות הראשית, היו קיימים אך נותרו ברובם ברקע ולא עוררו שינוי מהותי בהתפתחותו:

* *"She, in turn, seemed drawn to his ability to see patterns in seemingly random events."*

הקשר עם הדמויות המשניות, ובעיקר עם ורה (Vera) ,תואר כמשיכה רגשית ראשונית אך לא הוביל לעימות פנימי או שינוי בזהות.

* *"Elias had found an unexpected harmony here, a different kind of cosmic order."*

החיבור עם העובדים האחרים נשאר בעיקר כציר פונקציונלי, ולא גרם לו להתמודד עם זהותו לעומק.

* *"It was a world of exactitude, of measurements and mathematics—the antithesis of his fluid, interpretive craft. Yet Elias had found an unexpected harmony here, a different kind of cosmic order."*

כאן, מתואר הניגוד בין עולמו של אליאס לבין סביבת העבודה התעשייתית, אך לא נרמז שיש לכך השפעה עמוקה על דמותו או תהליך שינוי פנימי. אין עימות פנימי או שינוי ערכים שמובילים לעומק מסוים.

**בגרסה המשופרת**, חלק הדמויות תורם באופן מובהק להתפתחותו הפסיכולוגית והאישית של אליאס. לדוגמה:

* *"I've always lived between worlds," Elias confided to Vera... "Too scientific for the mystics, too mystical for the scientists."*

הקשרים מעמיקים ומאתגרים אותו להגדיר מחדש את זהותו – לא רק כאסטרולוג אלא כמגשר בין עולמות.

* *"Perhaps that's your strength," she replied... "We need bridges, not walls."*

ורה עצמה הופכת לדמות מפתח בהתמרה האישית של אליאס. דבריה מעצבים את ההבנה שלו את עצמו ואת ייעודו.

* *"Elias forwarded the emails to his personal account, his hands trembling. The stars had led him to this moment of choice..."*

גם הקשרים עם שאר הדמויות, למשל מרקו, עובדי הלילה, וריצ'רד גרייסון – אינם רק רקע לעלילה אלא זרזים להכרעות מוסריות שבונות את אישיותו.

**בגרסה המקורית, הדמויות שימשו בעיקר לתמיכה בתמה או לאווירה. לעומת זאת, בגרסה המשופרת הן ממלאות תפקיד מהותי בבניית הדמות הראשית. ההשפעה של ורה, של העובדים, ושל ריצ'רד היא ישירה, מעמיקה, ויוצרת שינוי אמיתי בזהות ובפעולה של אליאס. בכך, חלק זה עבר משימוש סמלי בלבד למנוע דרמטי מרכזי – והציון שופר בהתאם.**

**דוגמה לשיפור חלק "מלאכת הכתיבה" בעקבות ההמלצות האוטומטיות:**

**שם הסיפור : *The Uprising at Blackrock Mine***

**בסיפור זה נרשם שיפור בחלק מלאכת הכתיבה מציון של 0.8375 ל1 . להלן כמה**

**דוגמאות שיתארו את השיפור בחלק זה.**

**בגרסה המקורית,** סגנון הכתיבה היה אפקטיבי אך ניכר כי לעיתים נשען על תיאורים ישירים, מבנה ליניארי מדי וחזרתיות רעיונית. לדוגמה:

* *"Thomas pushed through the crowd. 'Enough! There will be no accusations without evidence. This settlement operates under law, not suspicion.'"*

המשפט אפקטיבי אך דיווחי , הוא "מספר" את רגשות הדמות והעלילה, מבלי להמחיש את עוצמת המתח באירוע.

* *"The mine remains closed until we've cleansed it of evil," Matthias announced.*

אמירה דרמטית, אך הדימויים נותרו שטחיים יחסית ולא נוצרת תחושת סכנה מוחשית בעזרת כתיבה סגנונית עשירה.

**בגרסה המשופרת,** מורגש שיפור משמעותי בתחום מלאכת הכתיבה: השפה עשירה יותר ומדויקת, תיאורי סביבה משתמשים בחושים מגוונים (ראייה, שמיעה, ריח), הדימויים חזקים ואטמוספריים והמבנה הקצבי יותר דרמטי ומתוזמן היטב. לדוגמה:

* *"The morning sun cast long shadows across the dusty settlement of Blackrock, a frontier mining town nestled between jagged mountains."*

פתיחה ויזואלית וחזקה המשלבת אור, צל, וסביבה באופן קולנועי.

* *"The acrid smell of sulfur and coal dust hung in the air... the constant drip of water... the distant echo of pickaxes... sounds that had become as familiar to him as his own heartbeat."*

תיאור רב־חושי (ריח, צליל, תחושה), שמייצר עומק סצנה נדיר ברמת כתיבה זו.

* *"Matthias smiled thinly. 'Not alone. I will accompany you... This is a spiritual matter.'"*

הדיאלוג מתוחכם ומובנה היטב, וכולל תת־טקסט ואיום מרומז.

* *"Inside the chamber, the ancient markings glowed with an unnatural blue light that cast eerie shadows across the rough-hewn walls."*

שימוש באור, צבע, ודימוי חזותי ליצירת מתח דרמטי עמוק ומדויק.

* *"The revelation transformed the tension in the room. Matthias's followers turned on him, their fear giving way to fury at being manipulated."*

שימוש דינמי בפעולה רגשית – מתיאור לתחושה, לשינוי רגשי, להשלכה דרמטית.

**המעבר מכתיבה פונקציונלית לכתיבה סוחפת, עתירת תיאורים ומבנה נרטיבי מודע לעצמו, בולט לכל אורכו של הסיפור. איכות הכתיבה המשופרת היא מהותית:**

* **תיאורי סביבה מורחבים יוצרים עולם חי ונושם.**
* **דיאלוגים מורכבים יותר, עם ניואנסים של רגש ותת-משמעות.**
* **שימוש מודע בקצב, מעברים חדים בין מתח לפעולה ולרגיעה.**

**לכן, העלייה בציון חלק "מלאכת הכתיבה" ניכרת וראויה. הסיפור עבר מגרסה אפקטיבית אך בסיסית לנראטיב עשיר, בוגר, ומשכנע.**

**ניתוח סטטיסטי לבדיקת שיפור הסיפורים:**

**שלב ראשון: בדיקת הנחות עבור מבחן T מזווג**

לפני שנפעיל את מבחן T מזווג (Paired T-Test), נבדוק קודם האם ההבדלים (הפרשים) בין ציוני "לפני" ו"אחרי" בכל תחום מתפלגים בצורה נורמלית.

לצורך כך השתמשנו ב־מבחן Shapiro-Wilk, שבודק האם ההתפלגות נורמלית:

|  |  |  |
| --- | --- | --- |
| **תחום** | **p-value (Shapiro-Wilk)** | **נורמליות מתקיימת?** |
| **דמויות** | **0.000013** | **לא נורמלי** |
| **קונפליקט** | **0.00000014** | **לא נורמלי** |
| **מלאכת הכתיבה** | **0.00012** | **לא נורמלי** |
| **היגיון** | **0.000000000000040** | **לא נורמלי** |
| **ציון כולל** | **0.000093** | **לא נורמלי** |

מסקנה: בכל התחומים, נשללה הנחת הנורמליות (p < 0.05).

לפיכך, **לא ניתן להפעיל את מבחן T מזווג**, ויש לעבור למבחן לא פרמטרי. **שלב שני: מעבר למבחן Wilcoxon**

בהמשך לכך ששללנו את האפשרות להשתמש במבחן T מזווג, נבצע מבחן Wilcoxon שמתאים לנתונים תלוים כאשר אין נורמליות.  
מבחן זה בודק האם קיים שינוי מובהק במדדים בין שני שלבים (לפני ואחרי) על בסיס הפרשי דירוגים.

**ניסוח השערות עבור כל תחום:**

**תחום- דמויות**

H₀: אין שינוי מובהק בציון התחום דמויות לפני ואחרי השיפור.

H₁: יש שינוי מובהק בציון התחום דמויות לאחר השיפור.

**תחום- קונפליקט**

H₀: אין שינוי מובהק בציון התחום קונפליקט לפני ואחרי השיפור.

H₁: יש שינוי מובהק בציון התחום קונפליקט לאחר השיפור.

**תחום- מלאכת הכתיבה**

H₀: אין שינוי מובהק בציון התחום מלאכת הכתיבה לפני ואחרי השיפור.

H₁: יש שינוי מובהק בציון התחום מלאכת הכתיבה לאחר השיפור.

**תחום- היגיון**

H₀: אין שינוי מובהק בציון התחום היגיון לפני ואחרי השיפור.

H₁: יש שינוי מובהק בציון התחום היגיון לאחר השיפור.

**תחום- ציון כולל**

H₀: אין שינוי מובהק בציון הכולל לפני ואחרי השיפור.

H₁: יש שינוי מובהק בציון הכולל לאחר השיפור.

**שלב שלישי: תוצאות מבחן Wilcoxon לכל תחום**

|  |  |  |  |
| --- | --- | --- | --- |
| **תחום הביקורת** | **T-Statistic** | **p-value** | **מובהקות** |
| **דמויות** | **323.5** | **8.73×10⁻⁸** | **מובהק** |
| **קונפליקט** | **402.5** | **5.28×10⁻⁵** | **מובהק** |
| **מלאכת הכתיבה** | **775.5** | **0.0222** | **מובהק** |
| **היגיון** | **145.0** | **1.63×10⁻⁷** | **מובהק** |
| **ציון כולל** | **988.0** | **9.73×10⁻⁷** | **מובהק** |

**שלב רביעי: מסקנות מההשוואה**

* בכלל התחומים שנבדקו – דמויות, קונפליקט, מלאכת הכתיבה, היגיון והציון הכולל – נצפתה עלייה ממוצעת בציונים לאחר השיפור.
* מאחר שהנחת הנורמליות לא התקיימה, הוחלט להשתמש ב־מבחן Wilcoxon Signed-Rank, המתאים לנתונים תלוים שאינם מתפלגים נורמלית.
* תוצאות המבחן הצביעו על כך ש־השיפור שנמדד בכל התחומים הוא מובהק סטטיסטית (p < 0.05).

**ניתוח ומסקנות:**

ממצאים שהתקבלו לאורך הניסוי מצביעים על כך שיישום ההמלצות שהופקו מהביקורת האוטומטית מביא לשיפור מדיד ומובהק במרבית תחומי ההערכה של הסיפורים. השיפור בא לידי ביטוי הן ברמה הכמותית – כפי שנמדד באמצעות ציונים כוללים וציונים נפרדים לכל תחום – והן ברמה האיכותית, כפי שהודגם בניתוחים מילוליים של שינויים במבנה הסיפור, הדמויות, הקונפליקט והכתיבה.

שימוש בשיטה סטטיסטית מתאימה (Wilcoxon Signed-Rank), לאחר שלילת הנחת הנורמליות, הוכיח באופן מובהק כי קיים שיפור מובהק סטטיסטית בכל התחומים שנבדקו.

באופן עקבי בשני האלגוריתמים שנבדקו Plotto ו־Plot Genie ,רוב הסיפורים הציגו עלייה בציון הכולל, תוך בולטות בשיפור במרכיבי ההיגיון והדמויות.

לסיכום, ניתן לקבוע כי המבקר האוטומטי מהווה כלי אפקטיבי לשיפור איכותי של סיפורים, הן בהיבט המבני והן בהיבט הסגנוני. השילוב בין הביקורת המובנית לבין הפקת שיפור ממוקד באמצעות מודלים שפתיים מתקדם ממלא את יעדיו, ומציע כיוון עתידי משמעותי לפיתוח מערכות תמיכה לכותבים ויוצרים בתחום הסיפורי.

**בנוסף לדו"ח זה, מצ"ב תיקיות הכוללת את כל הסיפורים שהופקו + דוחות ביקרות טיב הסיפור:**

1. Plotto – כולל את כל הסיפורים ודוחות הביקורת עבורן עבור סיפור שנוצרו באמצעות שלד שנוצר ע"י אלגוריתם Plotto.
2. Plot Genie– כולל את כל הסיפורים ודוחות הביקורת עבורן עבור סיפור שנוצרו באמצעות שלד שנוצר ע"י אלגוריתם Plot Genie.