# Manual Testing for Summarisation Accuracy – Text Insights

## Summary Table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Article Number | Article Name | Article Type | Acceptance Criteria Score | | |
| Short | Medium | Long |
| 1 | Skincare for Dry Skin | Casual | 11/20 | 13/20 | 15/20 |
| 2 | Arbitrage Betting | Casual | 12/20 | 17/20 | 17/20 |
| 3 | Cologne Description | Casual | 15/20 | 16/20 | 15/20 |
| 4 | Benefits of Sleep | Casual | 0/20 | 0/20 | 0/20 |
| 5 | Buying Guide for Kettles | Casual | 9/20 | 11/20 | 12/20 |
| 6 | Object-Oriented Programming | Academic | 14/20 | 11/20 | 9/20 |
| 7 | Porter’s Five Forces | Academic | 12/20 | 11/20 | 13/20 |
| 8 | Pythagoras’ Theorem | Academic | 16/20 | 16/20 | 17/20 |
| 9 | Newton’s First Law of Motion | Academic | 15/20 | 12/20 | 9/20 |
| 10 | Execution Architecture | Academic | 12/20 | 13/20 | 14/20 |

## How Acceptance Criteria is calculated

The Acceptance Criteria for this manual testing revolves around 2 scores: Legibility and Content. These scores are both a whole number between 0-10 and both scores are added evenly to make a total score for each Article out of 20.

Legibility is a score calculated on how easy-to-read each output is. This includes checking for grammatical and spelling errors while also considering the overall flow of the output. Finally, legibility will also consider any technical terms used in an output and how understandable they are in the context of the whole output.

Content, the second score in these tests, is a score that purely considers how well the output matches the content of the input. This score can dictate how well the output and summarization function is working in the overall context of each input.

Finally, in terms of how score should be read, a score of >=12 is generally ‘passable’, where the content is readable and has the right content, but may not have the necessary depth for learning. A score of >=16 means that the output did an exceptional job and clearly summarized the input in a meaningful way.

## Article 1: Skincare for Dry Skin

### Type of Article: Casual

### Complexity of Article: Complex

### Link to Article: <https://www.byrdie.com/dry-skincare-routine#:~:text=Dry%20Skin%20Ingredients%20to%20Look,shea%20butter%20and%20cocoa%20butter>.

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 8/10

Content: 3/10

Overall: 11/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 6/10

Overall: 13/20

### Long Summary Image:

### A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 8/10

Overall: 15/20

## Article 2: Arbitrage Betting

### Type of Article: Casual

### Complexity of Article: Simple

### Link to Article: <https://www.techopedia.com/gambling-guides/arbitrage-betting#:~:text=Arbitrage%20betting%20is%20to%20place,only%20occur%20at%20different%20sportsbooks>.

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 3/10

Overall: 12/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 8/10

Overall: 17/20

### Long Summary Image:

### A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 10/10

Overall: 17/20

## Article 3: Cologne Description

### Type of Article: Casual

### Complexity of Article: Simple

### Link to Article: <https://alexandriafragrances.com/products/himalaya-mountains?fbclid=IwAR0KIM7PJX7blwdztRfw03pRn_x49lx0jaCioK2XQG98G04ecLyzQeqHLIk>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 10/10

Content: 5/10

Overall: 15/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 7/10

Overall: 16/20

### Long Summary Image:

### A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 8/10

Content: 7/10

Overall: 15/20

## Article 4: Benefits of Sleep

### Type of Article: Casual

### Complexity of Article: Complex

### Link to Article: <https://newsinhealth.nih.gov/2013/04/benefits-slumber>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

0/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

0/20

### Long Summary Image:

### A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

0/20

## Article 5: Buying Guide for Kettles

### Type of Article: Casual

### Complexity of Article: Simple

### Link to Article: <https://www.reliancedigital.in/solutionbox/buying-guide-to-electric-kettles/>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 3/10

Overall: 9/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 5/10

Overall: 11/20

### Long Summary Image: A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 6/10

Overall: 12/20

## Article 6: Object Oriented Programming

### Type of Article: Academic

### Complexity of Article: Complex

### Link to Article: <https://www.techtarget.com/searchapparchitecture/definition/object-oriented-programming-OOP>

### Short Summary Image: A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 8/10

Content: 6/10

Overall: 14/20

### Medium Summary Image: A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 5/10

Overall: 11/20

### Long Summary Image:

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 5/10

Content: 4/10

Overall: 9/20

## Article 7: Porter’s Five Forces

### Type of Article: Academic

### Complexity of Article: Complex

### Link to Article: <https://www.mindtools.com/at7k8my/porter-s-five-forces>

### Short Summary Image:

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 5/10

Overall: 12/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 5/10

Content: 6/10

Overall: 11/20

### Long Summary Image: A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 7/10

Overall: 13/20

## Article 8: Pythagoras’ Theorem

### Type of Article: Academic

### Complexity of Article: Simple

### Link to Article: <https://en.wikipedia.org/wiki/Pythagorean_theorem>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 7/10

Overall: 16/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 7/10

Overall: 16/20

### Long Summary Image:

### A screenshot of a computer screen Description automatically generated

#### Acceptance Criteria Score:

Legibility: 9/10

Content: 8/10

Overall: 17/20

## Article 9: Newton’s First Law of Motion Explained

### Type of Article: Academic

### Complexity of Article: Complex

### Link to Article: <https://www.physicsclassroom.com/class/newtlaws/Lesson-1/Newton-s-First-Law>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 8/10

Content: 7/10

Overall: 15/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 5/10

Overall: 12/20

### Long Summary Image:

### A screenshot of a computer Description automatically generatedA screenshot of a computer screen Description automatically generated

#### Acceptance Criteria Score:

Legibility: 5/10

Content: 4/10

Overall: 9/20

## Article 10: Execution Architecture

### Type of Article: Academic

### Complexity of Article: Simple

### Link to Article: <https://esi.nl/research/output/methods/execution-architecture-analysis>

### Short Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 6/10

Overall: 12/20

### Medium Summary Image:

A screenshot of a computer

Description automatically generated

#### Acceptance Criteria Score:

Legibility: 6/10

Content: 7/10

Overall: 13/20

### Long Summary Image:

### A screenshot of a computer Description automatically generated

#### Acceptance Criteria Score:

Legibility: 7/10

Content: 7/10

Overall: 14/20

## Overall Findings

### Legibility:

Overall, the legibility of most outputs were quite understandable and followed proper English conventions. Overall, the legibility of a text was at its best at a ‘Default’ or ‘Long’ summary length, however there was an exception to this in Article 8: Newton’s First Law of Motion. In this article, as it contained mostly examples of the topic rather than information about the topic itself, the summary tried to combine too many elements of the article to try to explain the concept rather than treat most of it as an example. The overall recommendation is to keep most articles to ‘Default’ summary length.

### Content:

Overall, the content of most outputs varied immensely, with some articles like Article 2: Arbitrage Betting being an almost perfect summary of the topic, while Article 5: Buying Guide for Kettles, was almost useless in terms of its content. I believe this comes down to 2 factors that were defined within this testing process and have not been talked about until now.

### Academic Vs Casual Articles:

Articles thus far have been split into 2 categories, Academic and Casual. An Academic article can be seen as a more formal article that talks about some area of academic learning such as programming or mathematical concepts. Overall, the academic articles performed slightly worse than the casual articles and this may be attributed to the fact that there was either too much or too little content to explain to someone who is not well-versed in the topic. The time where the Academic summary performed well was on Article 8: Pythagoras’ Theorem. Having to explain only one formula and then being able to speak about the context around it seemed to be the perfect input for the summary to be able to consistently provide a meaningful output.

Casual articles can be seen as anything that is not an academic article; a more relaxed style of article that could potentially have more options for output because of a higher level of subjectivity. Surprisingly, this style of article performed better than its more objective counterpart, however the best summary by far was of a more objective topic, namely Article 2: Arbitrage Betting.

### Conclusion:

All in all, the summarization function works passably for now for most types of articles. Articles generally perform better if they are specific with what they are talking about and only a few examples of that topic are given. If a topic becomes too broad or too many examples are littered throughout the input, the summarization will usually fail in either legibility or content.