Flint Report

Project: Yamuder

Module: Authorship / Contract Cheating

File: Unknown doc.pdf

July 30, 2024

FLINT Systems, Inc.

info@flintai.com

+1 8337724946

flintai.com

Table of Contents

Table of Contents 2

1 Summary 3

2 Details Analysis 4

2.1 Feature with Positive Contributions 7

2.2 Feature with Negative Contributions 13

3 Documents Analyzed 18

3.1 Testing Document 18

3.2 Reference Documents 18

4 Feature Definitions 19

1 Summary

Based on the documents you uploaded for this analysis, our findings are within a 80% - 85% level of certainty that the submission was authored by **Robert Yamuder**, the reference author.

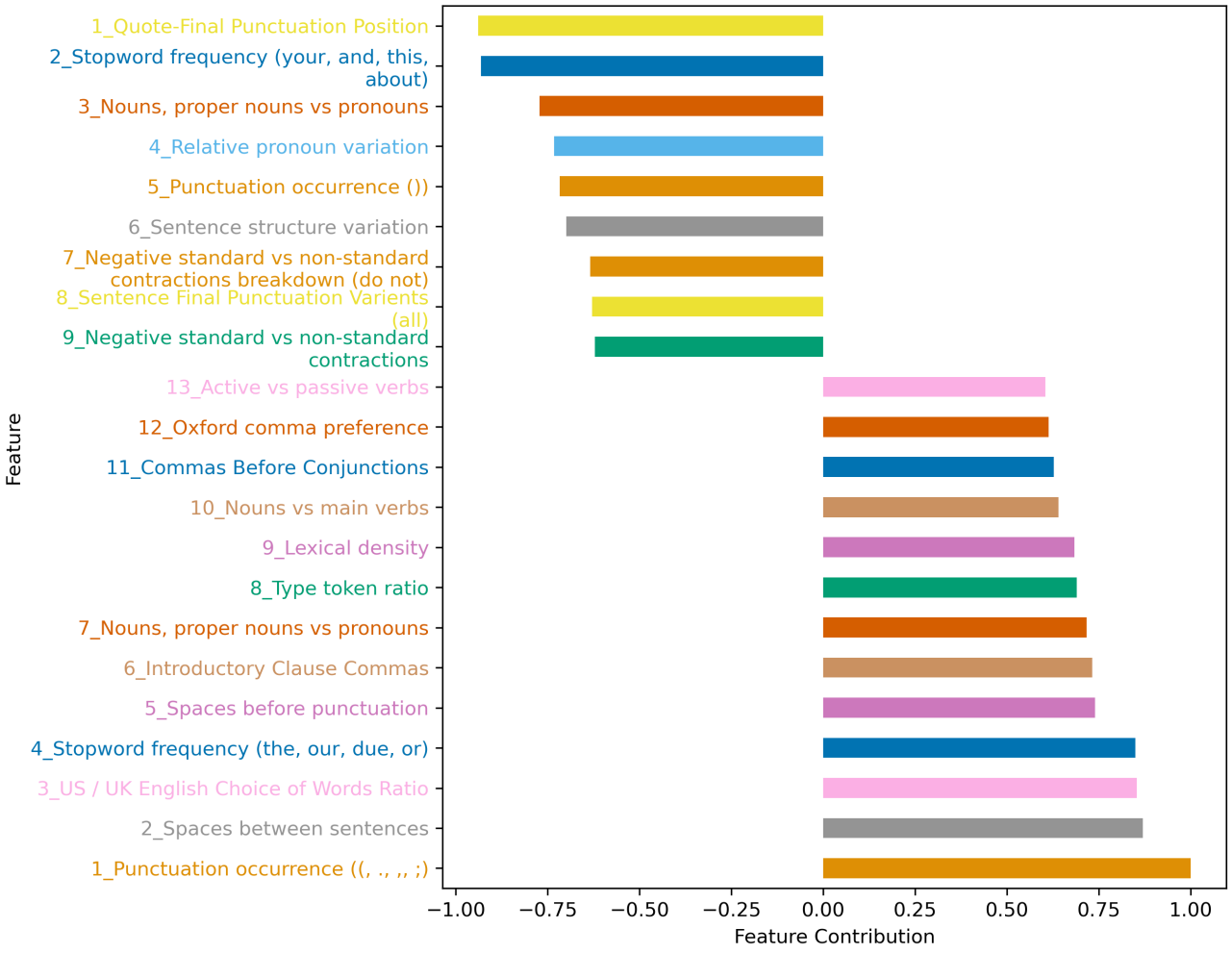
Report Key: As shown above, red color means the Q document is highly likely written by another author; yellow color means the Q document could be written by another/multiple authors; green color means the Q document is most likely written by **Robert Yamuder**, the author of reference.

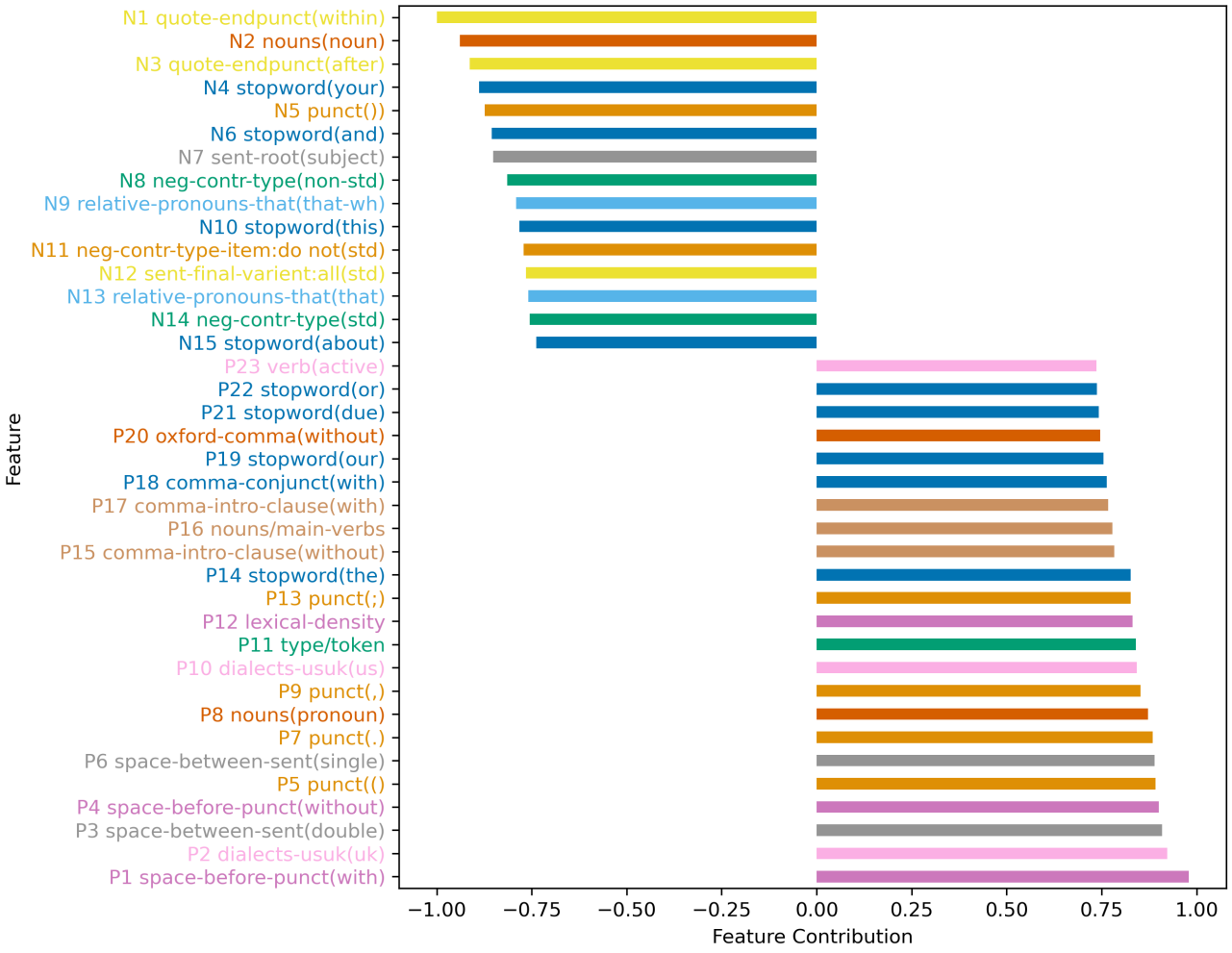
2 Details Analysis

The following report is generated for purpose of providing the results of an analysis of whether a certain document of questioned origin (*Q Document*) shares likely common authorship features with other documents of known origin (*K Documents*) applying standard linguistic rules and patterns.

The FLINT system performs its analysis by applying standard rules of statistical analysis when evaluating the documents and linguistic patterns used by the authors. The evaluated features include linguistic phenomena of various kinds, including grammatical patterns, punctuation and word usage. The more rare the feature, the more value that is associated with it, suggesting higher likelihood of unique document identification and commonality. As in any statistical analysis, the reported results are not conclusive. Rather the analysis reports the likelihood of a conclusion. Other facts and circumstances, beyond the scope of the FLINT system should be applied in order to reach a conclusive and final determination of fact.

The following table displays a stacked bar showing the ranked impact of the topmost linguistically important features that were applied in the current analysis, when evaluating the documents authored by the reference author Robert Yamuder compared to the Q Document. Features with positive contributions are those features suggests the compared documents are same author, while features with negative contributison are those features suggests the compared documents have different author.


1_Quote-Final Punctuation Position
2_Stopword frequency (your, and, this, about)
3_Nouns, proper nouns vs pronouns
4_Relative pronoun variation
5_Punctuation occurrence ())
6_Sentence structure variation
7_Negative standard vs non-standard contractions breakdown (do not)
8_Sentence Final Punctuation Varients (all)
9_Negative standard vs non-standard contractions
13_Active vs passive verbs
12_Oxford comma preference
11_Commas Before Conjunctions
10_Nouns vs main verbs
9_Lexical density
8_Type token ratio
7_Nouns, proper nouns vs pronouns
6_Introductory Clause Commas
5_Spaces before punctuation
4_Stopword frequency (the, our, due, or)
3_US / UK English Choice of Words Ratio
2_Spaces between sentences
1_Punctuation occurrence ((, ., ,, ;)
Feature
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Feature Contribution



Following is a detailed review of each of the features identified above.

2.1 Feature with Positive Contributions

1. Punctuation occurrence ((, ., ,, ;)

Frequency of different types of punctuation marks, such as ";", "!", ",".

P5 punct(()

Frequency of the punctuation mark "(" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	11.0	11
unknown	3.0	3

P7 punct(.)

Frequency of the punctuation mark "." in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	54.0	54
unknown	33.0	33

P9 punct(,)

Frequency of the punctuation mark "," in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	51.0	51
unknown	32.0	32

P13 punct(;)

Frequency of the punctuation mark ";" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	8.0	8
unknown	3.0	3

2. Spaces between sentences

Preference of the different numbers of empty spaces inserted between sentences.

P3 space-between-sent(double)

Ratio of sentences with double spaces between sentences compared to all instances of spaces between the sentences.

Author	Sample Mean Value	Matches
Robert Yamuder	0.08	4
unknown	0.27	9

P6 space-between-sent(single)

Ratio of sentences with a single space between sentences compared to all instances of spaces between the sentences.

Author	Sample Mean Value	Matches
Robert Yamuder	0.73	35
unknown	0.64	21

3. US / UK English Choice of Words Ratio

Preference of US vs. UK English words / spelling variants.

P2 dialects-usuk(uk)

Ratio of uk-specific words compared to all US/UK-specific words.

Author	Sample Mean Value	Matches
Robert Yamuder	0.33	1
unknown	0.0	0

P10 dialects-usuk(us)

Ratio of us-specific words compared to all US/UK-specific words.

Author	Sample Mean Value	Matches
Robert Yamuder	0.67	2
unknown	1.0	5

4. Stopword frequency (the, our, due, or)

Frequency of stopwords -- commonly used words which are filtered out in connection with the processing of natural language.

P14 stopword(the)

Frequency of the stopword "the" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	51.0	51
unknown	21.0	21

P19 stopword(our)

Frequency of the stopword "our" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	1
unknown	1.0	1

P21 stopword(due)

Frequency of the stopword "due" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	1
unknown	0.0	0

P22 stopword(or)

Frequency of the stopword "or" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	5.0	5
unknown	8.0	8

5. Spaces before punctuation

Preference of empty spaces before relevant punctuation punctuation compared to no empty spaces before relevant punctuation.

P1 space-before-punct(with)

Ratio of relevant punctuation with empty spaces before compared to all instances of relevant punctuation.

Author	Sample Mean Value	Matches
Robert Yamuder	0	0
unknown	0.0	0

P4 space-before-punct(without)

Ratio of relevant punctuation without empty spaces before compared to all instances of relevant punctuation.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	121
unknown	1.0	68

6. Introductory Clause Commas

Preferences for commas after sentence-initial phrases.

P15 comma-intro-clause(without)

Ratio of no commas after sentence-initial phrases as compared to all relevant sentence-initial phrases.

Author	Sample Mean Value	Matches
Robert Yamuder	0.43	9
unknown	0.33	3

P17 comma-intro-clause(with)

Ratio of commas after sentence-initial phrases as compared to all relevant sentence-initial phrases.

Author	Sample Mean Value	Matches
Robert Yamuder	0.57	12
unknown	0.67	6

7. Nouns, proper nouns vs pronouns

Ratio of pronouns (e.g., "I", "they"), common nouns (e.g., "shirt", "table"), and proper nouns (e.g., "Peter", "Jones").

P8 nouns(pronoun)

Ratio of pronouns compared to all types of nouns, proper nouns and pronouns.

Author	Sample Mean Value	Matches
Robert Yamuder	0.2	N/A
unknown	0.19	N/A

8. Type token ratio

Type/token ratio, i.e., the total number of unique words (types) divided by the total number of words (tokens).

P11 type/token

Author	Sample Mean Value	Matches
Robert Yamuder	44.18	N/A
unknown	50.27	N/A

9. Lexical density

Level of lexical density, i.e. ratio of nouns, verbs, adjectives, and adverbs compared to the total number of words.

P12 lexical-density

Author	Sample Mean Value	Matches
Robert Yamuder	40.19	N/A
unknown	45.86	N/A

10. Nouns vs main verbs

Ratio of nouns vs. verbs.

P16 nouns/main-verbs

Author	Sample Mean Value	Matches
Robert Yamuder	1.67	N/A
unknown	1.78	N/A

11. Commas Before Conjunctions

Preference for commas before conjunctions.

P18 comma-conjunct(with)

Ratio of commas before conjunctions as compared to all relevant conjunctions.

Author	Sample Mean Value	Matches
Robert Yamuder	0.06	1
unknown	0.19	5

12. Oxford comma preference

Preference of "Oxford commas" (commas inserted prior to the final conjunction in a serial list) (e.g. "one, two, and three") vs. non-insertion of "Oxford commas" in standard, structurally consistent, and non-ambiguous serial lists (e.g. "one, two and three").

P20 oxford-comma(without)

Ratio of standard serial lists without Oxford commas compared to all standard serial lists.

Author	Sample Mean Value	Matches
Robert Yamuder	0.5	0
unknown	0.5	2

13. Active vs passive verbs

Ratio of active (e.g., "eats") and passive verbs (e.g., "is eaten").

P23 verb(active)

Ratio of active verbs compared to all verbs.

Author	Sample Mean Value	Matches
Robert Yamuder	0.89	63
unknown	0.85	45

2.2 Feature with Negative Contributions

1. Quote-Final Punctuation Position

Preference of quote-final punctuation within or after a quote.

N1 quote-endpunct(within)

Ratio of quote-final punctuation within a quote as compared to all finally-punctuated quotes.

Author	Sample Mean Value	Matches
Robert Yamuder	0	0
unknown	1.0	1

N3 quote-endpunct(after)

Ratio of quote-final punctuation after a quote as compared to all finally-punctuated quotes.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	1
unknown	0.0	0

2. Stopword frequency (your, and, this, about)

Frequency of stopwords -- commonly used words which are filtered out in connection with the processing of natural language.

N4 stopword(your)

Frequency of the stopword "your" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	9.0	9
unknown	0.0	0

N6 stopword(and)

Frequency of the stopword "and" in the document.

Author	Sample Mean Value	Matches

Robert Yamuder	18.0	18
unknown	34.0	34

N10 stopword(this)

Frequency of the stopword "this" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	7.0	7
unknown	12.0	12

N15 stopword(about)

Frequency of the stopword "about" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	0	0
unknown	7.0	7

3. Nouns, proper nouns vs pronouns

Ratio of pronouns (e.g., "I", "they"), common nouns (e.g., "shirt", "table"), and proper nouns (e.g., "Peter", "Jones").

N2 nouns(noun)

Ratio of nouns compared to all types of nouns, proper nouns and pronouns.

Author	Sample Mean Value	Matches
Robert Yamuder	0.39	N/A
unknown	0.55	N/A

4. Relative pronoun variation

Preference of relative pronouns (e.g., "that") vs. "wh"-forms (e.g., "who", "which").

N9 relative-pronouns-that(that-wh)

Ratio of relative pronouns with "wh-" compared to all relative pronouns.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	6
unknown	0.2	1

N13 relative-pronouns-that(that)

Ratio of relative pronouns "that" compared to all relative pronouns.

Author	Sample Mean Value	Matches
Robert Yamuder	0	0
unknown	0.8	4

5. Punctuation occurrence ())

Frequency of different types of punctuation marks, such as ";", "!", ",".

N5 punct())

Frequency of the punctuation mark ")" in the document.

Author	Sample Mean Value	Matches
Robert Yamuder	17.0	17
unknown	3.0	3

6. Sentence structure variation

Sentence structure based on sentence-initial constituents (e.g., subject or object).

N7 sent-root(subject)

Ratio of sentences beginning with a subject compared to all sentences.

Author	Sample Mean Value	Matches
Robert Yamuder	0.41	18
unknown	0.7	23

7. Negative standard vs non-standard contractions breakdown (do not)

Preference of standard (e.g., "don't", "haven't") vs. non-standard negative contractions (e.g., "dont", "havent").

N11 neg-contr-type-item:do not(std)

Ratio of all standard negative constractions compared to all standard and non-standard negative contractions of "do not".

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	1
unknown	0.5	0

8. Sentence Final Punctuation Varients (all)

Preference of varients of sentence final punctuation orthography.

N12 sent-final-varient:all(std)

Ratio of standard punctuation compared with all punctuation used at the end of a sentence.

Author	Sample Mean Value	Matches
Robert Yamuder	1.0	1
unknown	0.5	0

9. Negative standard vs non-standard contractions

Preference of non-contracted (e.g., "is not", "does not") vs. contracted negative constructions (e.g., "isn't", "doesn’t").

N8 neg-contr-type(non-std)

Ratio of all non-standard negative constractions compared to all standard and non-standard negative contractions.

Author	Sample Mean Value	Matches
Robert Yamuder	0	0
unknown	0.5	0

N14 neg-contr-type(std)

Ratio of all standard negative constractions compared to all standard and non-standard negative contractions.

Author	Sample Mean Value	Matches

Robert Yamuder	1.0	1
unknown	0.5	0

3 Documents Analyzed

3.1 Testing Document

Following displays the document at question that was evaluated.

Name	Type	Word Count
Unknown doc.pdf	.pdf	1090

3.2 Reference Documents

Following displays the reference documents that were evaluated for comparison purposes.

Name	Type	Word Count
YamuderRef1.pdf	.pdf	1722

4 Feature Definitions

Feature	Description
Active vs passive verbs	Ratio of active (e.g., "eats") and passive verbs (e.g., "is eaten").
Commas Before Conjunctions	Preference for commas before conjunctions.
Introductory Clause Commas	Preferences for commas after sentence-initial phrases.
Lexical density	Level of lexical density, i.e. ratio of nouns, verbs, adjectives, and adverbs compared to the total number of words.
Negative standard vs non-standard contractions	Preference of non-contracted (e.g., "is not", "does not") vs. contracted negative constructions (e.g., "isn't", "doesn't").
Negative standard vs non-standard contractions breakdown	Preference of standard (e.g., "don't", "haven't") vs. non-standard negative contractions (e.g., "dont", "havent").
Nouns vs main verbs	Ratio of nouns vs. verbs.
Nouns, proper nouns vs pronouns	Ratio of pronouns (e.g., "I", "they"), common nouns (e.g., "shirt", "table"), and proper nouns (e.g., "Peter", "Jones").
Oxford comma preference	Preference of "Oxford commas" (commas inserted prior to the final conjunction in a serial list) (e.g. "one, two, and three") vs. non-insertion of "Oxford commas" in standard, structurally consistent, and non-ambiguous serial lists (e.g. "one, two and three").
Punctuation occurrence	Frequency of different types of punctuation marks, such as ";", "!", ",".
Quote-Final Punctuation Position	Preference of quote-final punctuation within or after a quote.
Relative pronoun variation	Preference of relative pronouns (e.g., "that") vs. "wh"-forms (e.g., "who", "which").
Sentence Final Punctuation Varients	Preference of varients of sentence final punctuation orthography.
Sentence structure variation	Sentence structure based on sentence-initial constituents (e.g., subject or object).
Spaces before punctuation	Preference of empty spaces before relevant punctuation punctuation compared to no empty spaces before relevant punctuation.
Spaces between sentences	Preference of the different numbers of empty spaces inserted between sentences.
Stopword frequency	Frequency of stopwords -- commonly used words which are filtered out in connection with the processing of natural language.
Type token ratio	Type/token ratio, i.e., the total number of unique words (types) divided by the total number of words (tokens).

US / UK English Choice of Words Ratio	Preference of US vs. UK English words / spelling variants.