A Modern Appraisal:

What Machine Learning Exposes About Clausewitz & Jomini

MAJ Thomas P. Malejko

30 March 2020

Western civilization loves a good rivalry, so much so that societies will create an antagonistic relationship—whether one truly existed or not—just to debate the subject: John Locke v. Thomas Hobbes; Alexander Hamilton v. Aaron Burr; Thomas Edison v. Nikola Tesla; and, of course, Carl von Clausewitz v. Antoine-Henri Jomini. However, these assessments of intellectual polarity are the subject of human opinion and, therefore, bias and fallacy. What would remain of the rivalry if the human was removed from the assessment? How truly different are the intellectual rivals? This is where history and mathematics—another diametric relationship—merge to definitively analyze the prominent writings of the aforementioned renowned military theorists. Confirming what many historians have suspected for decades, modern analysis—using machine learning and natural language processing—concludes that Clausewitz and Jomini wrote about war at fundamentally different levels. While Clausewitz's *On War*[1] broadly evaluated war and the national implications thereof, Jomini's *The Art of War*[2] largely centered on the practical applications of warfare.

The technical approach used during this comparative analysis is a typical data science workflow, where each historical document's text was imported, cleaned, explored, and analyzed. During the first two phases of the aforementioned data science process, each book was reformatted and normalized so all that remained of each document was its substantive and consequential text. The exploratory data analysis of this text, which combined with relevant domain knowledge, provided initial insights that informed the approaches used during the two advanced analyses. Topic modelling, the first advanced analysis, is an unsupervised machine learning method that discovers latent—or hidden—topics in a series of documents. This analysis generated multiple probabilistic topic models from the combined texts of *On War* and *The Art of War*, where the best model demonstrated an ideal blend of good quantitative metrics (such as

high topic coherence) and qualitative sense. Next, the best model evaluated every chapter in *On War* and *The Art of War*, resulting in a machine-predicted topic for each chapter. Annex A (Topic Modelling Workflow) provides a more thorough explanation of this process that includes step-by-step examples. The second advanced technique used during this analysis was association rule mining, which is an unsupervised machine learning method that identifies relationships between variables—in this case, words. To derive maximum insights from this computationally expensive analysis, word associations were only generated for sentences that contain the keywords 'war' and 'nation.' Simply, association rule mining will reveal words that commonly and frequently appear in sentences with each keyword.

The initial exploratory analysis of the two documents revealed fundamental differences between the principal writings of Clausewitz and Jomini—differences that would persist throughout the entire analysis. Visualized in two wordclouds (Figure 1, below), the words used by each author exposed fundamentally different views of the same topic.

Figure 1. Wordclouds depicting the consequential words used by Clausewitz (left) and Jomini (right) in On War *and* The Art of War, *respectively. The size of the word corresponds to its frequency of use in each text—that is the larger the word appears in the visualization, the more times it was used.*

In *On War*, Clausewitz used the term 'war' most frequently followed by terms that demonstrated his complete appreciation for the subject—from the tactical to the strategic and the theoretical to the practical. Conversely, Jomini used 'army' twice as frequently as war highlighting his preference to write about the pragmatic nature of conflict. While Clausewitz frequently used strategic and political terms such as 'case,' 'force(s),' 'object,' and 'mean(s)'—all of which are top 15 words in *On War*—'force(s)' is the only strategic-political term to appear in Jomini's most frequently used 20 words. Instead, Jomini's word usage reveals an effort to reduce war to its tactical elements supported by practical examples as evidenced by a frequent use of 'line(s),' 'operation,' 'point,' and 'thousand.'[3] Jomini did use the term 'strategic' often; however, his usage of the term referred to all manner of war between politics and combat, which may confuse contemporary observers that associate strategy with politics.[4] These findings confirm long-held, historical assessments about the writings of the two authors such as those held by Dr. Christopher Bassford who wrote that Clausewitz's *On War* was a "ruthless examination of military-theoretical ideas" while Jomini tended to be "utilitarian and…didactic."[5]

Topic modelling reinforces the concept that Clausewitz's *On War* evaluated the concept of war more broadly than Jomini's *The Art of War*. Jointly, the two authors wrote about five topics (as identified by machine learning): 1) Tactics; 2) Army Organization and Sustainment; 3) Practical Applications and Historical Examples; 4) Theory of Warfare; and 5) The Character of War. Machine-predicted topics for each chapter reveal that *On War* addressed all five topics while *The Art of War* largely centered on a single topic—Practical Applications and Historical Examples. It is important to note that *The Art of War* likely includes aspects of all topics, however the preponderance of text in each chapter—more often than not—is about the practical application of warfare, which ultimately determined the results seen in Figure 2 (below).
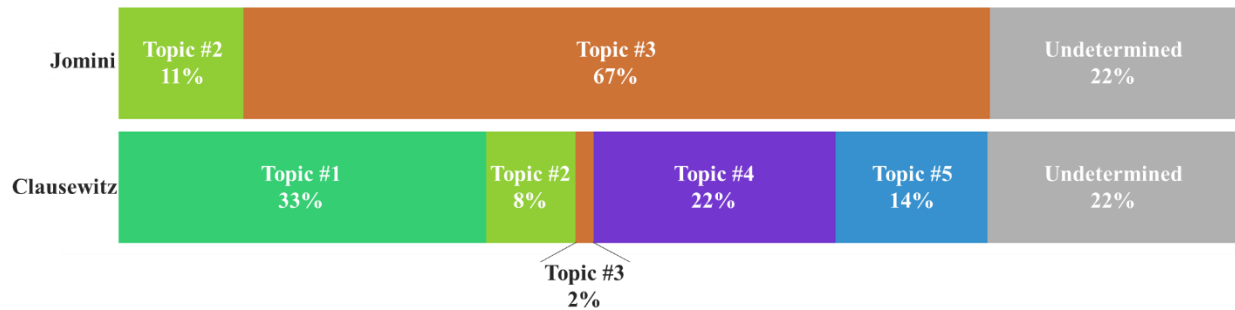
Figure 2. The percent of chapters in On War *and* The Art of War *assigned a given topic during topic modelling, an unsupervised machine learning method.*[6]

Consider the format of *The Art of War's* opening chapter, wherein Jomini repeatedly states an 'article' of war then supports that principle with a historical example. The frequent use of vignettes is very practical, especially for an author wanting "to convince, [and] to establish the way that men think about war,"[7] but skews the amount of text written about each topic towards the 'Practical Applications and Historical Examples' category. The results of the topic modelling analysis also reinforce a historical conception about *On War*; that its frequent abridgement typically obscures the true breadth of Clausewitz's work.[8] While the initial books—and their associated chapters—of *On War* largely focus on the theory and character of war, the text's latter elements (Books 6 and 7) are very tactically oriented.[9] These results quantitatively confirm what historians, such as Mr. John Shy, have long-said—that Clausewitz largely wrote to satisfy his understanding of war, while Jomini wrote to publish and impress.[10]

Finally, association rule mining evaluated with network analysis revealed that Clausewitz and Jomini wrote about 'war' and the 'nation' through fundamentally different lenses. While 'war' was—unsurprisingly—the central theme of sentences that contained the keyword 'war,' differences emerged in terms of breadth. While Jomini largely discussed 'war' in the context of practical or historical matters, Clausewitz demonstrated a deeper understanding of the subject. In the two network diagrams (Figure 3, below), notice that Clausewitz frequently commented on the purpose and theory of warfare—concepts not found in Jomini's equivalent network diagram.

Consider the first (1) and second (2) clusters in Clausewitz's diagram, wherein his purpose of war can be roughly observed—that is the object of war is to exert influence over or result in the destruction of one's enemy. Similarly, the third (3) cluster shows that Clausewitz often discusses the theory and conduct of war throughout his work. Jomini—by contrast—does not frequently discuss either topic in *The Art of War*, instead exclusively focusing on pragmatic elements, including his well-known line(s) and zone(s).
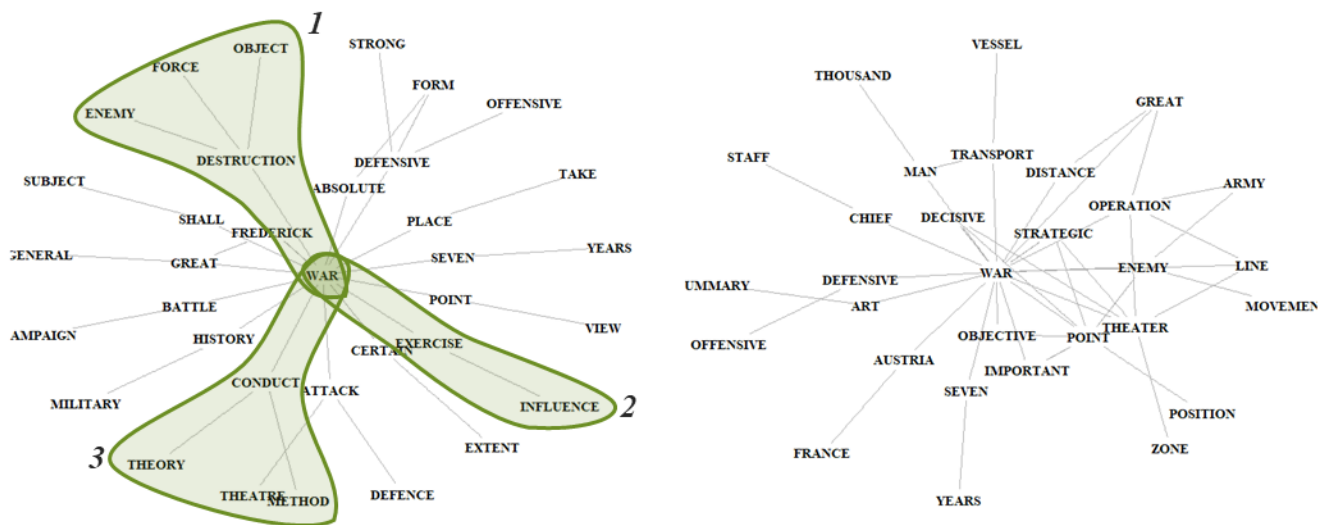

Figure 3. Network diagrams for On War *(left) and* The Art of War *(right) depicting words that frequently appear in conjunction with the keyword 'war' with significant community clusters highlighted to show the key differences between the two diagrams.*

In a similar analysis of sentences containing the keyword 'nation,' a suprising difference emerged between the two authors. While 'nation' was the central theme in Clausewitz's sentences on the subject, Jomini generally referred to the 'nation' only in the context of its 'army' and 'force(s).' Consider the following two sentences as prototypical examples of Jomini's use of the keyword 'nation' in *The Art of War*:

> 1. We are far from thinking that any purely military maxims can insure *(sic)* the success of remote invasions: in four thousand years only five or six have been successful, and in a hundred instances they have nearly ruined <u>nations</u> and <u>armies</u>.[11]

> 2. No <u>army</u>, however disciplined, can contend successfully against such a system applied to a great <u>nation</u>, unless it be strong enough to hold all the essential points of the country,

cover its communications, and at the same time furnish an active <u>force</u> sufficient to beat the enemy wherever he may present himself.[12]

In both instances, Jomini only used the term 'nation' in the context of an army and in practical matters—in the first, as a description of a broad historical observation and, in the second, as a lesson learned from Napoleon's peninsular campaign. These examples, combined with the network diagram in Figure 4 (below right), reveal that Jomini viewed the 'nation' as just another element of war.
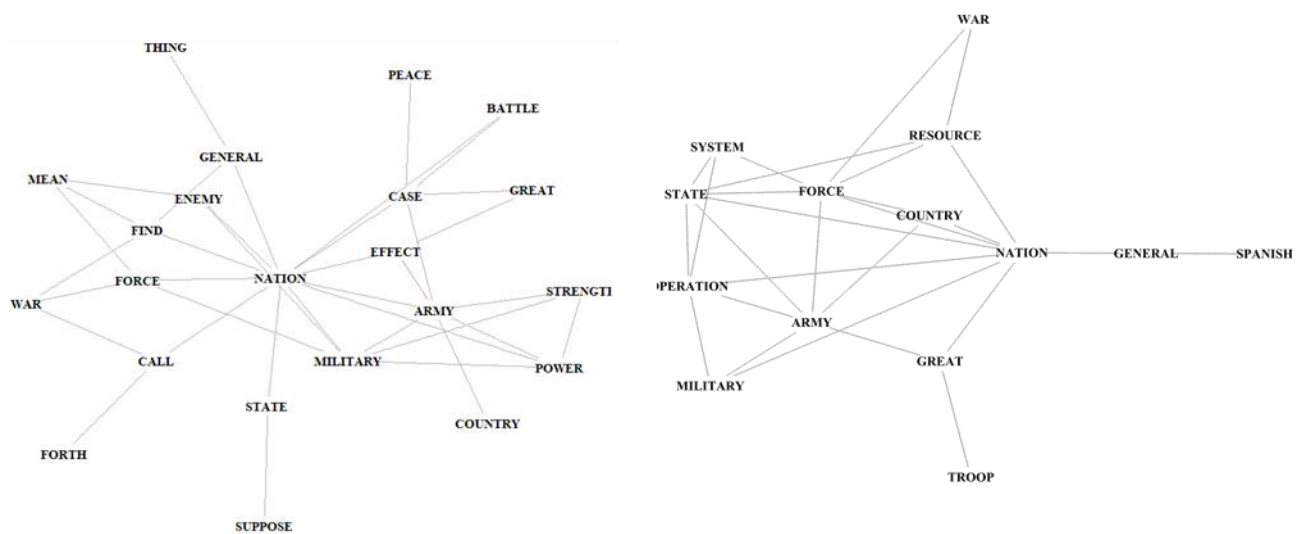

Figure 4. Network diagrams for On War *(left) and* The Art of War *(right) depicting words that frequently appear in sentences alongside the keyword 'nation.' Notice that 'nation' occupies a central position in the left diagram, while—on the right—the term 'nation' is not.*

Clausewitz, by contrast, viewed the 'nation' as the nexus of warfare, hence the term's central position in *On War's* network diagram (above left). In summary, both association rules analyses demonstrate that Clausewitz and Jomini wrote about warfare—from the nation to the act itself—from fundamentally different viewpoints.

Advanced analytics, including machine learning and natural language processing, definitively reinforce traditional views about the differences between the two military theorists. Jomini's *The Art of War* is a pragmatic look at warfare that uses practical and historical vignettes to emphasize the tactical employment of an army. On the other hand, Clausewitz's *On War* is a

broad evaluation of the title subject—from the tactical employment of an army to the tangible

interaction of politics and war to ethereal characterizations of combat itself. These dissimilar

approaches to writing about the same topic demonstrate that Clausewitz and Jomini not only

possessed different purposes for authorship but diverging views on war.

Annex A (Topic Modelling Workflow)

The principal output of a machine generated topic model is a list of words associated with each topic, which analysts must summarize for ready interpretability. While the table below only shows the top 10 words associated with each topic, the assigned topic title was based on a review of the top 30 words.

Id	Topic Names	Top 10 Words Associated with Topic
1	Tactics	*mountain, defence, defender, river, assailant, defensive, attack, position, fortress, offensive*
2	Army Organization & Sustainment	*division, advanced, march, guard, corps, column, cantonment, wing, cavalry, mile*
3	Practical Applications & Historical Examples	*napoleon, operation, center, maneuver, defense, thousand, officer, surprise, zone, staff*
4	Theory of Warfare	*theory, combat, destruction, strategy, political, activity, mind, knowledge, act, motive*
5	Character of War	*victory, battle, conqueror, gravity, loss, decision, centre, combat, theatre, conquer*

Table A.1. The five machine learning topics generated from On War *and* The Art of War.

The following table shows a complete topic modelling workflow for a few sentences from each document to highlight the process of text normalization and demonstrate topic prediction based on that single sentence. Note, that as part of normalization text is: case normalized; lemmatized (converted to its root—or lemma—word); and removed if it is a common term (such as a pronoun, conjunction, or article).

Author	Raw Text	Normalized Text	Predicted Topic
Clausewitz	"…it is made in single marches with a longer or shorter interval between them, and at each halting point the assailant is just as much acting on the defensive as his adversary."[13]	single, marche, long, short, interval, halting, point, assailant, act, defensive, adversary	Tactics
Jomini	"There are large corps dispatched to a distance from the zone of	large, corps, dispatch, distance, zone,	Army Organization,

	operations of the main army, in order to make diversions of greater or less importance."[14]	operation, main, army, order, diversion, great, importance	Sustainment, & the Troops
Clausewitz	"We take first Buonaparte in his famous enterprise against Blücher's Army in February 1814…"[15]	bonaparte, famous, enterprise, blücher, army, february	Practical Applications & Historical Examples
Jomini	"If the theory of war leaves any points unprovided for, that of retreats is certainly one of them."[16]	theory, war, leave, point, unprovide, retreat, certainly	Theory of War
Clausewitz	"Each strives by physical force to compel the other to submit to his will…"[17]	strive, physical, force, compel, submit	Character of War

Figure A.2. Demonstration of text normalization and topic assignment process for sample texts from On War *and* The Art of War.

End Notes

1. Carl von Clausewitz, On War, trans. J. J. Graham (Gutenberg Project, 2021), accessed February 1, 2021, https://www.gutenberg.org/files/1946/1946-0.txt.

2. Antoine Henri Jomini, The Art of War, trans. Wm. P. Craighill and George H. Mendell (Gutenberg Project, 2004), accessed February 1, 2021, https://www.gutenberg.org/cache/epub/13549/pg13549.txt.

3. This edition of The Art of War used in this analysis included numbers written as words, such as 'two hundred thousand.' Consequently, Jomini's common description of troop sizes in his vignettes meant that 'thousand' became a high frequency word in this text.

4. John Shy, "Jomini" in *Makers of Modern Strategy from Machiavelli to the Nuclear Age*, ed. Peter Paret with Gordon A. Craig and Felix Gilbert (Princeton, New Jersey: Princeton University Press, 1986), 167-168.

5. Christopher Bassford, "Jomini and Clausewitz—Their Interaction," Clausewitz.Com, last modified February 26, 1993, accessed February 21, 2021, http://www.clausewitz.com/readings/Bassford/Jomini/JOMINIX.htm.

6. A chapter assigned a "undetermined" indicates that a topic could not be conclusively predicted (defined as having a prediction probability in excess of .45).

7. John Shy, "Jomini," 163.

8. Bassford, "Jomini and Clausewitz—Their Interaction."

9. In On War, the topic modelling algorithm predicted that 71 percent of chapters in Books 6 and 7 were about tactics, whereas not a single chapter in the text's first four chapters received that topic prediction.

10. John Shy, "Jomini," 158.

11. Jomini, *The Art of War*, Line 6042.

12. Ibid, Line 732.

13. Clausewitz, *On War*, Line 14,147.

14. Jomini, *The Art of War*, Line 7797.

15. Clausewitz, *On War*, Line 6522.

16. Jomini, *The Art of War*, Line 8313.

17. Clausewitz, *On War*, Line 26.

Bibliography

Books

Howard, Michael. "The Influence of Clausewitz*."* In *On War,* Carl von Clausewitz, 27-44.
      Princeton, New Jersey: Princeton University Press, 1989.

Paret, Peter. "The Genesis of *On War."* In *On War,* Carl von Clausewitz, 3-25. Princeton, New
      Jersey: Princeton University Press, 1989.

Shy, John. "Jomini." In *Makers of Modern Strategy from Machiavelli to the Nuclear Age*, edited
      by Peter Paret with Gordon A. Craig and Felix Gilbert, 143-185. Princeton, New Jersey:
      Princeton University Press, 1986.

Electronic Sources

Bassford, Christopher. "Jomini and Clausewitz--Their Interaction." *Clausewitz.Com*. Last
      modified February 26, 1993. Accessed February 21, 2021.
      http://www.clausewitz.com/readings/Bassford/Jomini/JOMINIX.htm.

Clausewitz, Carl von. *On War*. Translated by J. J. Graham. Gutenberg Project, 2021. Accessed
      February 1, 2021. https://www.gutenberg.org/files/1946/1946-0.txt.

Caidid. "Machine Learning Text Analysis of *On War*." *The Children's Illustrated Clausewitz*,
      December 8, 2017. Accessed February 12, 2021. https://clausewitzforkids.com/.

Jomini, Antoine Henri. *The Art of War*. Translated by Wm. P. Craighill and George H. Mendell.
      Gutenberg Project, 2004. Accessed February 1, 2021.
      https://www.gutenberg.org/cache/epub/13549/pg13549.txt.

Technical Sources (Inspiration for Association Rule Mining Analysis)

Gates, Ami M. "Association Rule Mining with Tweets: Thinking Outside the Basket" Presented
      at the DC R Conference, Washington D.C., March 26, 2019. Accessed July 8, 2020.
      https://www.youtube.com/watch?v=eOOhn9CX2qU.