Grade Key: Program #5 Machine Learning and the Federalist Papers

Produce a professionally formatted Jupyter Notebook that shows an experimental workflow to generate evidence for the Disputed Federalist Papers.

Submit
  (1) a README.txt,
  (2) a Jupyter Notebook   [no, some, yes, ++]   _____

Writing
  Markdown notebook looks professional   [no, some, yes, ++]   _____

  Summary sets the context   [no, some, yes, ++]   _____

  Mention/cite Jockers and Witten's 2010 paper;   [no, some, yes, ++]   _____
    mention *their* results on the disputed papers   [no, some, yes, ++]   _____

  References included   [no, some, yes, ++]   _____

Methods (when you run different techniques, you may need to change details for each run)

  Mention starting data set (e.g., 1- or 2-grams, from *Lexos*)   [no, some, yes, ++]   _____
      (You need to explain where the data came from,
          and how tokenized)

  Detecting "bad data"
  -   missing features (column values); any change to words used? [no, some, yes, ++] _____
  -   Provide an argument why for removing any outliers or not.   [no, some, yes, ++] _____

  Discuss cluster and/or classification algorithms applied   [no, some, yes, ++]   _____

Results

  Generate Results and discuss evidence for the authorship of disputed papers:

     (a) a clustering technique   [no, some, yes, ++]   _____

     (b) a classification technique   [no, some, yes, ++]   _____

     (c) an additional classification technique   [no, some, yes, ++]   _____

For each category, the assessment rank of [no, some, yes, ++] will indicate the following associated level of completion:

Level of completion	Effort
No:	None
Some:	Average
Yes:	Above average
++:	Superior