abelardo g fournier 2 abelardo\_g\_fournier\_2

# README

This is a script that estimates the similarity between documents that are taken two by two from a larger collection. The algorithm calculates the average frequencies of the words used within the collection of all the documents, and compares these values with the local frequencies of the words within each document individually. The deviations fluctuations become then the parameters used to calculate a quantitative value of semantic proximity.

### More information:

http://scikit-learn.org/stable/modules/feature\_extraction.html#text-feature-extraction (http://scikitlearn.org/stable/modules/feature\_extraction.html#text-feature-extraction)

```
FEATURE EXTRACTION: MOST FREQUENT WORDS
An_Ethnography_of_Error.txt
-- error(13) human(12) problem(11) driverless(9) car(9)
Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- hamlet(22) hes(20) data(16) war(12) vietnam(12)
Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- space(21) data(18) cables(15) metadata(14) tor(12)
From_Page_Rank_to_Rankbrain.txt
-- search(28) google(22) page(17) users(13) algorithms(11)
Machine_Listening.txt
-- human(28) wavenet(22) ernst(18) machine(16) listening(16)
Machine_Pedagogies.txt
-- images(16) human(15) freire(15) oppressor(13) learning(13)
Participation_in_Infrastructures.txt
-- systems(20) electricity(19) demand(13) time(12) resources(11)
Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- data(26) recognition(11) digital(11) capture(11) perception(9)
Relearn_to_Read_Speed_Readers.txt
-- reading(39) speed(35) reader(13) readers(12) new(12)
Resolution_Theory.txt
-- resolution(66) term(17) word(15) resolutions(15) formal(14)
Testing_Texting_South_a_Political_Fiction.txt
-- language(18) south(17) political(17) fictions(13) 2015(11)
The_Cultural_Politics_of_Information_and_of_Debt.txt
-- information(20) debt(19) communication(19) terranova(15) message(14)
The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- language(14) communication(14) reality(11) thought(10) medium(10)
The_Stupid_Network_that_we_Know_and_Love.txt
-- network(21) ip(15) tcp(13) internet(13) protocol(12)
```

Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt -- vision(13) media(11) land(10) earth(10) practices(7)

```
CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS
              -- 0.099 The Stupid Metwork_that_we.Know_and_Love.txt
-- 0.099 The Stupid_Metwork_that_we.Know_and_Love.txt
-- 0.0497 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.046 Computing_Mor_Morratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.045 Resolution_Theory.txt
-- 0.045 Resolution_Theory.txt
          {\tt Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt}
        Computing_Nar_Narratives_The_Namlet_Evaluation_System_in_Yietman.txt
-- 0.881 Testing_Texting_South_o_Political_Fiction_txt
-- 0.878 Pattern=Recognition-across=Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.879 Participation_in_Infrastructures.txt
-- 0.879 Participation_in_Infrastructures.txt
-- 0.865 From_Page_Rank_to_Rankbrain.txt
-- 0.895 Testypid_Metwork_that_we_Know_and_Love.txt
-- 0.895 Testypid_Metwork_that_we_Know_and_Love.txt
-- 0.897 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.897 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.897 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.895 Teach_in_In_Edoc_Dead_Readers_txt
-- 0.896 An_Ethnography_of_Error_txt
-- 0.896 Machine_Pedagagies_txt
-- 0.896 Machine_Listering_txt
-- 0.893 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.893 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.802 Resolution_Theory_txt
        Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
                       0.106 Participation_in_Infrastructures.txt
0.105 The_Stupid_Network_that_we_Know_and_Love.txt
                       0.103 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt 0.092 The_Cultural_Politics_of_Information_and_of_Debt.txt
 - 0.092 The Cultural_Politics_of_Information_and_of_Debt.txt
- 0.084 The Signification_Communication_Question_Some_Initial_Remarks.txt
- 0.0875 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
- 0.0606 From_Page_Rank_to_Rankbrain.txt
- 0.0657 Computing_Mar_Narratives_The_Manlet_Evaluation_System_in_Vietman.txt
- 0.055 Relearn_to_Read_Speed_Readers.txt
- 0.055 Relearn_to_Read_Speed_Readers.txt
- 0.054 Machine_Pedagogies.txt
- 0.052 Machine_Listening_txt
- 0.054 Resolution_Theory_txt
- 0.047 An_Ethnography_of_Error.txt
  From_Page_Rank_to_Rankbrain.txt

- 0.138 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.091 The_Stupid_Network_that_we_Know_and_Love.txt

- 0.080 Rolearn_to_Read_Speed_Readers.txt

- 0.080 Relearn_to_Read_Speed_Readers.txt

- 0.071 Testing_Texting_South_a_Political_Fiction.txt

- 0.068 Machine_Listening.txt

- 0.068 Machine_Listening.txt
             -- 0.068 Machine_Listening.txt
-- 0.067 An_Ethnography.of_Error.txt
-- 0.065 Computing_Mar_Marratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.060 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.057 Participation_in_Infrastructures.txt
-- 0.054 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.054 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.054 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.048 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.040 Resolution_Theory.txt
Machine_Listening.txt
-- 0.119 Machine_Pedagogies.txt
-- 0.115 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.185 Testing_Texting_South_a_Political_Fiction.txt
-- 0.085 An_Ethnography_of_Error.txt
-- 0.085 An_Ethnography_of_Error.txt
-- 0.085 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.083 Relearn_to_Read_Speed_Readers.txt
-- 0.070 Time_Stupid_Network_that_we_Know_and_Love.txt
                    0.069 The_Cultural_Politics_of_Information_and_of_Debt.txt
0.068 From_Page_Rank_to_Rankbrain.txt
                    0.058 Participation_in_Infrastructures.txt
0.052 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
          -- 0.050 Resolution_Theory.txt
-- 0.036 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
 Machine_Pedagogies.txt

- 0.19 Machine_Listening.txt

- 0.088 An_Ethnography_of_Error.txt

- 0.087 From_Page_Rank_to_Rankbrain.txt

- 0.088 Partern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.087 Dentern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.073 Unmaking_Screens_a_Genealogy_of_the_Mineral_vision.txt

- 0.073 Unmaking_Screens_a_Genealogy_of_the_Mineral_vision.txt

- 0.074 The_Signification_Communication_Question_Some_Initial_Remarks.txt

- 0.059 The_Stupid_Network_that_we_Know_and_Love.txt

- 0.059 The_Stupid_Network_that_we_Know_and_Love.txt

- 0.054 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt

- 0.049 Participation_in_Infrastructures.txt

- 0.049 Computing_Nor_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt

- 0.036 Resolution_Theory.txt
-- 0.036 Resolution_Theory.txt

Participation_in_Infrastructures.txt
-- 0.108 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.108 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.106 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.084 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.080 Pattern_Recognition-across_Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.072 Testing_Texting_South_a_Political_Fiction.txt
-- 0.072 Relearn_to_Read_Speed_Readers.txt
-- 0.072 Relearn_to_Read_Speed_Readers.txt
-- 0.073 An_Ethongraphy_of_Error.tx
-- 0.074 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.063 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.065 Resolution_Theory.txt
-- 0.065 Resolution_Theory.txt
-- 0.057 From_Page_Rank_to_Rankbrain.txt
-- 0.045 Machine_Istening_txt
     Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
  -- 0.138 From_Page_Rank_to_Rankbrain.txt
-- 0.103 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.097 Relearn_to_Read_Speed_Readers_txt
-- 0.096 Testing_Texting_South_a_Political_Fiction.txt
-- 0.096 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.084 Machine_Listening.txt
-- 0.088 Participation_in_Infrastructures.txt
-- 0.080 Participation_in_Infrastructures.txt
-- 0.078 Computing_Nor_Norratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.078 Computing_Nor_Norratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.078 Computing_Nor_Norratives_The_Mineral_Vision.txt
-- 0.078 Computing_Nor_Norratives_The_Mon_and_Love.txt
-- 0.060 Resolution_Theory.txt
-- 0.060 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.061 An_Ethnography_of_Error.txt
          -- 0.138 From_Page_Rank_to_Rankbrain.txt
```

0.125 Testing\_Texting\_South\_a\_Political\_Fiction.txt
0.097 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

0.083 Machine\_Listening.txt
0.082 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt 0.080 From\_Page\_Rank\_to\_Rankbrain.txt

- 0.074 Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

- 0.063 The Stupid.Network\_that\_we\_Know\_and\_Love.txt
- 0.060 An\_Ethnography\_of\_Error\_txt
- 0.060 An\_Ethnography\_of\_Error\_txt
- 0.057 Computing\_Mor\_Narrotives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt
- 0.056 Machine\_Pedagogies.txt
- 0.055 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt
- 0.052 The\_Culturul\_Politics\_of\_Information\_and\_of\_Debt.txt
- 0.043 Resolution\_Theory.txt

0.069 Pattern-Recognition-coross-Bodies-and-Machines-by-Anarchival-Means.txt
0.067 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt
0.063 Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

0.047 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt 0.045 An\_Ethnography\_of\_Error.txt 0.044 The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt

0.040 From\_Page\_Rank\_to\_Rankbrain.txt 0.038 Testing\_Texting\_South\_a\_Political\_Fiction.txt 0.036 Machine\_Pedagogies.txt
0.024 Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt

Testing\_Texting\_South\_a\_Political\_Fiction.txt

- 0.134 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

- 0.125 Relearn\_to\_Read\_Speed\_Readers.txt

- 0.096 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.093 Ummaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

- 0.086 Machine\_Listening.txt

0 081 Computing War Narratives The Hamlet Evaluation System in Vietman txt

0.071 From\_Page\_Rank\_to\_Rankbrain.txt

0.066 An\_Ethnography\_of\_Error.txt
0.061 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt

0.055 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

-- 0.038 Resolution\_Theory.txt

-- 0.092 Elusive\_Borders\_Aesthetic\_Perpectualization\_Ine\_Space\_lime\_of\_Metadata.i
-- 0.069 Machine\_Istening.txt
-- 0.068 Machine\_Pedagogies.txt
-- 0.068 Machine\_Pedagogies.txt
-- 0.067 Resolution\_Ineory.txt
-- 0.067 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.063 Participation\_In\_Infrastructures.txt
-- 0.061 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.061 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.054 Texn\_Page\_Rank\_ty\_Bankbrain\_txt
-- 0.054 Texn\_Page\_Rank\_ty\_Bankbrain\_txt
-- 0.055 Relearn\_to\_Read\_Speed\_Readers.txt
-- 0.052 Relearn\_to\_Read\_Speed\_Readers.txt

- 0.044 An\_Ethnography\_of\_Error.txt - 0.034 Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt

0.164 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt
 0.134 Testing\_Texting\_South\_a\_Political\_Fiction.txt

0.085 The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt
0.084 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

## 

## RECOMMENDED READINGS FLOW

If you enjoyed reading An\_Ethnography\_of\_Error.txt, you might like also:

-- Machine\_Pedagogies.txt

-- Machine\_Listening.txt

If you enjoyed reading Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt, you might like also:

-- Testing\_Texting\_South\_a\_Political\_Fiction.txt

-- Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

If you enjoyed reading Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt, you might like also:

-- Participation\_in\_Infrastructures.txt

-- The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt

If you enjoyed reading From\_Page\_Rank\_to\_Rankbrain.txt, you might like also:

-- Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

-- The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt

If you enjoyed reading Machine\_Listening.txt, you might like also:

-- Machine\_Pedagogies.txt

-- The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

If you enjoyed reading Machine\_Pedagogies.txt, you might like also:

-- Machine\_Listening.txt

-- An\_Ethnography\_of\_Error.txt

If you enjoyed reading Participation\_in\_Infrastructures.txt, you might like also:

-- Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

-- The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt

If you enjoyed reading Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt, you might like also:

-- From\_Page\_Rank\_to\_Rankbrain.txt

-- Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

If you enjoyed reading Relearn\_to\_Read\_Speed\_Readers.txt, you might like also:

-- Testing\_Texting\_South\_a\_Political\_Fiction.txt

-- Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

If you enjoyed reading Resolution\_Theory.txt, you might like also:

-- The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

-- Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt

If you enjoyed reading Testing\_Texting\_South\_a\_Political\_Fiction.txt, you might like also:

-- The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

-- Relearn\_to\_Read\_Speed\_Readers.txt

If you enjoyed reading The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt, you might like also:

-- The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

-- The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt

If you enjoyed reading The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt, you might like also:

-- The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt

-- Testing\_Texting\_South\_a\_Political\_Fiction.txt

If you enjoyed reading The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt, you might like also: -- Participation\_in\_Infrastructures.txt

-- Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

If you enjoyed reading Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt, you might like also:

-- Participation\_in\_Infrastructures.txt

-- Testing\_Texting\_South\_a\_Political\_Fiction.txt

-- 0.066 Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt
-- 0.048 From\_Page\_Rank\_to\_Rankbrain.txt
-- 0.033 Computing\_War\_Marratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt

-- 0.033 Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt

The\_Stupid\_Network\_that\_we\_Know\_and\_Love.txt
-- 0.107 Participation\_in\_Infrastructures.txt
-- 0.105 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt
-- 0.093 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt
-- 0.093 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt
-- 0.085 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt
-- 0.085 Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt
-- 0.087 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.078 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.078 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.079 Machine\_Listening.txt
-- 0.063 Relearn\_to\_Read\_Speed\_Readers.txt
-- 0.063 Relearn\_to\_Read\_Speed\_Readers.txt
-- 0.059 Machine\_Pedagagies.txt
-- 0.059 Machine\_Pedagagies.txt
-- 0.059 Machine\_Pedagagies.txt
-- 0.059 Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt
-- 0.044 Resolution\_Theory.txt

Unmaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

-- 0.108 Participation\_in\_Infrastructures.txt

-- 0.82 The\_Stupid\_Network\_that\_we\_Know\_ond\_Love.txt

-- 0.082 The\_Stupid\_Network\_that\_we\_Know\_ond\_Love.txt

-- 0.077 Pattern-Recognition\_arcross-Bodies-and-Machines-by-Anarchival-Means.txt

-- 0.076 Machine\_Listening.txt

-- 0.075 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

-- 0.076 Relearn\_to\_Road\_Speed\_Readers.txt

-- 0.074 An\_Ethnography\_of\_Error.txt

-- 0.073 Machine\_Pedagogies.txt

-- 0.079 Computing\_War\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt

-- 0.066 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

-- 0.061 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt

-- 0.063 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt