

# 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://scikit-learn.org/stable/modules/feature\_extraction.html#text-feature-extraction) (Proximus NV → Level 3 Communications, Inc. → GitHub, Inc.)

//// 1 / 3 //////////////////////////////////////

## 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)

```
//// 3 / 3 //////////////////////////////////////
CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS
An_Ethnography_of_Error.txt
-- 0.004 Machine_Listening.txt
-- 0.074 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.073 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.067 From_Page_Rank_to_Rankbrain.txt
-- 0.066 Testing_Texting_South_a_Political_Fiction.txt
-- 0.065 Participation_in_Infrastructures.txt
-- 0.066 Returns_to_Read_Speed_Readers.txt
-- 0.070 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.047 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.046 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.045 Resolution_Theory.txt
-- 0.046 The_Cultural_Politics_of_Information_and_of_Debt.txt
Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.061 Testing_Texting_South_a_Political_Fiction.txt
-- 0.070 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.070 Participation_in_Infrastructures.txt
-- 0.065 From_Page_Rank_to_Rankbrain.txt
-- 0.070 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.067 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.067 Returns_to_Read_Speed_Readers.txt
-- 0.066 An_Ethnography_of_Error.txt
-- 0.066 Machine_Listening.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.065 The_Signification_Communication_Question_Some_Initial_Remarks.txt
Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.100 Participation_in_Infrastructures.txt
-- 0.100 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.100 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.092 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.090 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.075 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.068 From_Page_Rank_to_Rankbrain.txt
-- 0.057 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.055 Returns_to_Read_Speed_Readers.txt
-- 0.055 Testing_Texting_South_a_Political_Fiction.txt
-- 0.054 Machine_Listening.txt
-- 0.052 Resolution_Theory.txt
-- 0.047 An_Ethnography_of_Error.txt
From_Page_Rank_to_Rankbrain.txt
-- 0.140 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.087 Machine_Listening.txt
-- 0.086 Returns_to_Read_Speed_Readers.txt
-- 0.072 Testing_Texting_South_a_Political_Fiction.txt
-- 0.067 An_Ethnography_of_Error.txt
-- 0.066 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.066 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.067 Participation_in_Infrastructures.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.066 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.064 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.060 Resolution_Theory.txt
Machine_Listening.txt
-- 0.110 Machine_Listening.txt
-- 0.110 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.086 Testing_Texting_South_a_Political_Fiction.txt
-- 0.075 An_Ethnography_of_Error.txt
-- 0.084 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.076 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.070 The_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
-- 0.065 Participation_in_Infrastructures.txt
-- 0.052 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.046 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
Machine_Pedagogies.txt
-- 0.110 Machine_Listening.txt
-- 0.086 An_Ethnography_of_Error.txt
-- 0.086 From_Page_Rank_to_Rankbrain.txt
-- 0.080 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.072 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.066 Returns_to_Read_Speed_Readers.txt
-- 0.066 An_Ethnography_of_Error.txt
-- 0.064 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.065 Participation_in_Infrastructures.txt
-- 0.046 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.036 Resolution_Theory.txt
Participation_in_Infrastructures.txt
-- 0.100 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.107 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.090 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 Returns_to_Read_Speed_Readers.txt
-- 0.070 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.065 An_Ethnography_of_Error.txt
-- 0.064 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.065 Resolution_Theory.txt
-- 0.065 Machine_Listening.txt
-- 0.067 From_Page_Rank_to_Rankbrain.txt
Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.140 From_Page_Rank_to_Rankbrain.txt
-- 0.107 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.097 Returns_to_Read_Speed_Readers.txt
-- 0.096 Testing_Texting_South_a_Political_Fiction.txt
-- 0.090 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.084 Machine_Listening.txt
-- 0.080 Participation_in_Infrastructures.txt
-- 0.080 Machine_Pedagogies.txt
-- 0.076 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.077 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.066 Resolution_Theory.txt
-- 0.066 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.061 An_Ethnography_of_Error.txt
Returns_to_Read_Speed_Readers.txt
-- 0.110 Testing_Texting_South_a_Political_Fiction.txt
-- 0.097 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.093 Machine_Listening.txt
-- 0.086 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.080 From_Page_Rank_to_Rankbrain.txt
-- 0.077 Participation_in_Infrastructures.txt
-- 0.065 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.067 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.066 An_Ethnography_of_Error.txt
-- 0.066 Machine_Pedagogies.txt
-- 0.067 From_Page_Rank_to_Rankbrain.txt
Resolution_Theory.txt
-- 0.110 Testing_Texting_South_a_Political_Fiction.txt
-- 0.093 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.089 Machine_Listening.txt
-- 0.086 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.080 From_Page_Rank_to_Rankbrain.txt
-- 0.077 Participation_in_Infrastructures.txt
-- 0.065 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.061 An_Ethnography_of_Error.txt
Testing_Texting_South_a_Political_Fiction.txt
-- 0.110 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.110 Returns_to_Read_Speed_Readers.txt
-- 0.093 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.093 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.086 Machine_Listening.txt
-- 0.086 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.075 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.077 Participation_in_Infrastructures.txt
-- 0.077 From_Page_Rank_to_Rankbrain.txt
-- 0.065 An_Ethnography_of_Error.txt
-- 0.066 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.065 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.066 Machine_Pedagogies.txt
-- 0.058 Resolution_Theory.txt
The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.100 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.093 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.090 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.089 Machine_Listening.txt
-- 0.087 Resolution_Theory.txt
-- 0.080 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.081 Participation_in_Infrastructures.txt
-- 0.081 Testing_Texting_South_a_Political_Fiction.txt
-- 0.081 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.080 From_Page_Rank_to_Rankbrain.txt
-- 0.082 Returns_to_Read_Speed_Readers.txt
-- 0.084 An_Ethnography_of_Error.txt
-- 0.034 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.100 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.110 Testing_Texting_South_a_Political_Fiction.txt
-- 0.110 Machine_Listening.txt
-- 0.080 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.080 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.080 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.084 Participation_in_Infrastructures.txt
-- 0.080 Returns_to_Read_Speed_Readers.txt
-- 0.075 An_Ethnography_of_Error.txt
-- 0.075 Machine_Pedagogies.txt
-- 0.075 Resolution_Theory.txt
-- 0.080 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.085 From_Page_Rank_to_Rankbrain.txt
-- 0.085 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
-- 0.100 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.093 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.092 From_Page_Rank_to_Rankbrain.txt
-- 0.090 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.092 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.070 Testing_Texting_South_a_Political_Fiction.txt
-- 0.070 Machine_Listening.txt
-- 0.066 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.063 Returns_to_Read_Speed_Readers.txt
-- 0.060 Machine_Pedagogies.txt
-- 0.059 An_Ethnography_of_Error.txt
-- 0.059 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.046 Resolution_Theory.txt
Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.100 Testing_Texting_South_a_Political_Fiction.txt
-- 0.091 Testing_Texting_South_a_Political_Fiction.txt
-- 0.086 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.077 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.076 Machine_Listening.txt
-- 0.075 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.075 Returns_to_Read_Speed_Readers.txt
-- 0.074 An_Ethnography_of_Error.txt
-- 0.074 Machine_Pedagogies.txt
-- 0.070 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.066 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.063 Resolution_Theory.txt
-- 0.061 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.054 From_Page_Rank_to_Rankbrain.txt
```

//// 3 / 3 //////////////////////////////////////

## 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

