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)

```
//// 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)
```

```
//// 2 / 3 ////////////////////////////////////////////////////////////////////////

CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS

An_Ethnography_of_Error.txt
-- 0.088 Machine_Pedagogies.txt
-- 0.085 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
-- 0.062 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.060 Relearn_to_Read_Speed_Readers.txt
-- 0.059 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
-- 0.044 The_Cultural_Politics_of_Information_and_of_Debt.txt

Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.081 Testing_Texting_South_a_Political_Fiction.txt
-- 0.078 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.070 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.070 Participation_in_Infrastructures.txt
-- 0.065 From_Page_Rank_to_Rankbrain.txt
-- 0.059 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.058 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.057 Relearn_to_Read_Speed_Readers.txt
-- 0.046 An_Ethnography_of_Error.txt
-- 0.040 Machine_Pedagogies.txt
-- 0.036 Machine_Listening.txt
-- 0.034 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.033 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.024 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
-- 0.088 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.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 Relearn_to_Read_Speed_Readers.txt
-- 0.055 Testing_Texting_South_a_Political_Fiction.txt
-- 0.054 Machine_Pedagogies.txt
-- 0.050 Machine_Listening.txt
-- 0.047 Resolution_Theory.txt
-- 0.047 An_Ethnography_of_Error.txt

From_Page_Rank_to_Rankbrain.txt
-- 0.118 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.091 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.088 Participation_in_Infrastructures.txt
-- 0.080 Relearn_to_Read_Speed_Readers.txt
-- 0.075 Testing_Texting_South_a_Political_Fiction.txt
-- 0.068 Machine_Listening.txt
-- 0.067 An_Ethnography_of_Error.txt
-- 0.065 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.060 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.058 Machine_Pedagogies.txt
-- 0.054 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.052 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.044 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
-- 0.105 Testing_Texting_South_a_Political_Fiction.txt
-- 0.085 An_Ethnography_of_Error.txt
-- 0.084 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.083 Relearn_to_Read_Speed_Readers.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
-- 0.058 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
-- 0.119 Machine_Listening.txt
-- 0.088 An_Ethnography_of_Error.txt
-- 0.087 From_Page_Rank_to_Rankbrain.txt
-- 0.080 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.073 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.071 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.068 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.066 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.056 Relearn_to_Read_Speed_Readers.txt
-- 0.054 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.049 Testing_Texting_South_a_Political_Fiction.txt
-- 0.045 Participation_in_Infrastructures.txt
-- 0.040 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.036 Resolution_Theory.txt

Participation_in_Infrastructures.txt
-- 0.108 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.107 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.106 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.094 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.088 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
-- 0.070 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.065 An_Ethnography_of_Error.txt
-- 0.063 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.061 Resolution_Theory.txt
-- 0.058 Machine_Listening.txt
-- 0.057 From_Page_Rank_to_Rankbrain.txt
-- 0.045 Machine_Pedagogies.txt

Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.118 From_Page_Rank_to_Rankbrain.txt
-- 0.101 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
-- 0.092 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.088 Participation_in_Infrastructures.txt
-- 0.088 Machine_Pedagogies.txt
-- 0.078 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.077 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.069 Resolution_Theory.txt
-- 0.066 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.064 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.061 An_Ethnography_of_Error.txt

Relearn_to_Read_Speed_Readers.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
-- 0.076 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.072 Participation_in_Infrastructures.txt
-- 0.063 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.060 An_Ethnography_of_Error.txt
-- 0.057 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.056 Machine_Pedagogies.txt
-- 0.055 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.054 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.043 Resolution_Theory.txt

Resolution_Theory.txt
-- 0.071 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.069 Pattern-Recognition-across-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
-- 0.061 Participation_in_Infrastructures.txt
-- 0.058 Machine_Listening.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
-- 0.043 Relearn_to_Read_Speed_Readers.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
-- 0.123 Relearn_to_Read_Speed_Readers.txt
-- 0.096 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
-- 0.081 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.078 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.072 Participation_in_Infrastructures.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
-- 0.050 Machine_Pedagogies.txt
-- 0.038 Resolution_Theory.txt

The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.164 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.093 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.092 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.088 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.068 Machine_Pedagogies.txt
-- 0.067 Resolution_Theory.txt
-- 0.064 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 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.054 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

The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.164 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.136 Testing_Texting_South_a_Political_Fiction.txt
-- 0.115 Machine_Listening.txt
-- 0.094 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.093 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.084 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.084 Participation_in_Infrastructures.txt
-- 0.082 Relearn_to_Read_Speed_Readers.txt
-- 0.073 An_Ethnography_of_Error.txt
-- 0.071 Machine_Pedagogies.txt
-- 0.071 Resolution_Theory.txt
-- 0.066 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.048 From_Page_Rank_to_Rankbrain.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
-- 0.091 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.091 From_Page_Rank_to_Rankbrain.txt
-- 0.085 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.082 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.078 Testing_Texting_South_a_Political_Fiction.txt
-- 0.070 Machine_Listening.txt
-- 0.069 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.063 Relearn_to_Read_Speed_Readers.txt
-- 0.059 An_Ethnography_of_Error.txt
-- 0.059 Machine_Pedagogies.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
-- 0.093 Testing_Texting_South_a_Political_Fiction.txt
-- 0.082 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
-- 0.074 Relearn_to_Read_Speed_Readers.txt
-- 0.074 An_Ethnography_of_Error.txt
-- 0.071 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
```

abelardo_g_fournier_2

```
//// 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
```