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)

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