

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)

```
//// 2 / 3 //////////////////////////////////////
CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS

An_Ethnography_of_Error.txt
-- 0.888 Machine_Pedagogies.txt
-- 0.885 Machine_Listening.txt
-- 0.874 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.873 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.867 From_Page_Rank_to_Rankbrain.txt
-- 0.866 Testing_Texting_South_a_Political_Fiction.txt
-- 0.865 Participation_in_Infrastructures.txt
-- 0.861 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.860 Relearn_to_Read_Speed_Readers.txt
-- 0.859 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.847 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.846 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.845 Resolution_Theory.txt
-- 0.844 The_Cultural_Politics_of_Information_and_of_Debt.txt

Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt
-- 0.881 Testing_Texting_South_a_Political_Fiction.txt
-- 0.878 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.870 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.870 Participation_in_Infrastructures.txt
-- 0.869 From_Page_Rank_to_Rankbrain.txt
-- 0.869 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.857 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.857 Relearn_to_Read_Speed_Readers.txt
-- 0.846 An_Ethnography_of_Error.txt
-- 0.840 Machine_Pedagogies.txt
-- 0.836 Machine_Listening.txt
-- 0.834 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.833 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.834 Resolution_Theory.txt

Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.186 Participation_in_Infrastructures.txt
-- 0.188 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.189 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.882 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.884 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.875 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.869 From_Page_Rank_to_Rankbrain.txt
-- 0.857 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.855 Relearn_to_Read_Speed_Readers.txt
-- 0.855 Testing_Texting_South_a_Political_Fiction.txt
-- 0.854 Machine_Pedagogies.txt
-- 0.852 Machine_Listening.txt
-- 0.847 Resolution_Theory.txt
-- 0.847 An_Ethnography_of_Error.txt

From_Page_Rank_to_Rankbrain.txt
-- 0.138 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.881 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.887 Machine_Pedagogies.txt
-- 0.880 Relearn_to_Read_Speed_Readers.txt
-- 0.871 Testing_Texting_South_a_Political_Fiction.txt
-- 0.867 Machine_Listening.txt
-- 0.867 An_Ethnography_of_Error.txt
-- 0.865 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.868 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.857 Participation_in_Infrastructures.txt
-- 0.854 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.854 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.848 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.848 Resolution_Theory.txt

Machine_Listening.txt
-- 0.119 Machine_Pedagogies.txt
-- 0.115 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.886 Testing_Texting_South_a_Political_Fiction.txt
-- 0.885 An_Ethnography_of_Error.txt
-- 0.884 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.883 Relearn_to_Read_Speed_Readers.txt
-- 0.880 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.869 From_Page_Rank_to_Rankbrain.txt
-- 0.868 Participation_in_Infrastructures.txt
-- 0.852 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.850 Resolution_Theory.txt
-- 0.836 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt

Machine_Pedagogies.txt
-- 0.119 Machine_Listening.txt
-- 0.888 An_Ethnography_of_Error.txt
-- 0.887 From_Page_Rank_to_Rankbrain.txt
-- 0.880 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.871 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.872 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.868 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.859 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.856 Relearn_to_Read_Speed_Readers.txt
-- 0.851 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.849 Testing_Texting_South_a_Political_Fiction.txt
-- 0.845 Participation_in_Infrastructures.txt
-- 0.848 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.836 Resolution_Theory.txt

Participation_in_Infrastructures.txt
-- 0.188 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.187 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.187 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.884 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.886 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.872 Testing_Texting_South_a_Political_Fiction.txt
-- 0.872 Relearn_to_Read_Speed_Readers.txt
-- 0.870 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.865 An_Ethnography_of_Error.txt
-- 0.863 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.859 Machine_Listening.txt
-- 0.857 From_Page_Rank_to_Rankbrain.txt
-- 0.845 Machine_Pedagogies.txt

Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.138 From_Page_Rank_to_Rankbrain.txt
-- 0.183 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.887 Relearn_to_Read_Speed_Readers.txt
-- 0.886 Testing_Texting_South_a_Political_Fiction.txt
-- 0.882 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.884 Machine_Listening.txt
-- 0.880 Participation_in_Infrastructures.txt
-- 0.880 Machine_Pedagogies.txt
-- 0.878 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.877 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.869 Resolution_Theory.txt
-- 0.866 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.864 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.861 An_Ethnography_of_Error.txt

Relearn_to_Read_Speed_Readers.txt
-- 0.125 Testing_Texting_South_a_Political_Fiction.txt
-- 0.897 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.897 Machine_Listening.txt
-- 0.882 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.880 From_Page_Rank_to_Rankbrain.txt
-- 0.874 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.872 Participation_in_Infrastructures.txt
-- 0.863 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.860 An_Ethnography_of_Error.txt
-- 0.857 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.856 Machine_Pedagogies.txt
-- 0.851 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.847 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.843 Resolution_Theory.txt

The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.872 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.869 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.861 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.863 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.861 Participation_in_Infrastructures.txt
-- 0.859 Machine_Listening.txt
-- 0.847 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.845 An_Ethnography_of_Error.txt
-- 0.844 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.843 Relearn_to_Read_Speed_Readers.txt
-- 0.840 From_Page_Rank_to_Rankbrain.txt
-- 0.838 Testing_Texting_South_a_Political_Fiction.txt
-- 0.836 Machine_Pedagogies.txt
-- 0.834 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt

Testing_Texting_South_a_Political_Fiction.txt
-- 0.114 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.125 Relearn_to_Read_Speed_Readers.txt
-- 0.889 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.893 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.896 Machine_Listening.txt
-- 0.881 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.878 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.872 Participation_in_Infrastructures.txt
-- 0.871 From_Page_Rank_to_Rankbrain.txt
-- 0.866 An_Ethnography_of_Error.txt
-- 0.861 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.853 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.849 Machine_Pedagogies.txt
-- 0.838 Resolution_Theory.txt

The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.164 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.883 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.882 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.880 Machine_Listening.txt
-- 0.880 Machine_Pedagogies.txt
-- 0.867 Resolution_Theory.txt
-- 0.864 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.863 Participation_in_Infrastructures.txt
-- 0.861 Testing_Texting_South_a_Political_Fiction.txt
-- 0.861 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.861 From_Page_Rank_to_Rankbrain.txt
-- 0.852 Relearn_to_Read_Speed_Readers.txt
-- 0.844 An_Ethnography_of_Error.txt
-- 0.834 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt

The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.164 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.114 Testing_Texting_South_a_Political_Fiction.txt
-- 0.115 Machine_Listening.txt
-- 0.891 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.885 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.884 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.884 Participation_in_Infrastructures.txt
-- 0.881 Relearn_to_Read_Speed_Readers.txt
-- 0.873 An_Ethnography_of_Error.txt
-- 0.871 Machine_Pedagogies.txt
-- 0.871 Resolution_Theory.txt
-- 0.866 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.864 From_Page_Rank_to_Rankbrain.txt
-- 0.833 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt

The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.187 Participation_in_Infrastructures.txt
-- 0.185 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.881 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.881 From_Page_Rank_to_Rankbrain.txt
-- 0.885 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.882 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.878 Testing_Texting_South_a_Political_Fiction.txt
-- 0.870 Machine_Listening.txt
-- 0.866 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.863 Relearn_to_Read_Speed_Readers.txt
-- 0.859 An_Ethnography_of_Error.txt
-- 0.859 Machine_Pedagogies.txt
-- 0.859 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.844 Resolution_Theory.txt

Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.188 Participation_in_Infrastructures.txt
-- 0.881 Testing_Texting_South_a_Political_Fiction.txt
-- 0.881 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.877 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.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

-- 0.876 Machine_Listening.txt
-- 0.875 Elusive_Borders_Aesthetic_Perpetualization_The_Space_Time_of_Metadata.txt
-- 0.874 Relearn_to_Read_Speed_Readers.txt
-- 0.874 An_Ethnography_of_Error.txt
-- 0.873 Machine_Pedagogies.txt
-- 0.870 Computing_for_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.866 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.863 Revolution_Theory.txt
-- 0.861 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.854 From_Page_Rank_to_RankBrain.txt