

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_Vietnam.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 //////////////////////////////////////

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_Vietnam.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

//// 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_Vietnam.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
-- 0.865 From_Page_Rank_to_Rankbrain.txt
-- 0.859 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
-- 0.824 Resolution_Theory.txt

Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.186 Participation_in_Infrastructures.txt
-- 0.185 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.183 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.182 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.184 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.175 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.169 From_Page_Rank_to_Rankbrain.txt
-- 0.167 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.165 Relearn_to_Read_Speed_Readers.txt
-- 0.155 Testing_Texting_South_a_Political_Fiction.txt
-- 0.154 Machine_Pedagogies.txt
-- 0.152 Resolution_Theory.txt
-- 0.147 An_Ethnography_of_Error.txt

From_Page_Rank_to_Rankbrain.txt
-- 0.138 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.131 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.127 Machine_Pedagogies.txt
-- 0.126 Relearn_to_Read_Speed_Readers.txt
-- 0.121 Testing_Texting_South_a_Political_Fiction.txt
-- 0.120 Machine_Listening.txt
-- 0.117 An_Ethnography_of_Error.txt
-- 0.116 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.115 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.114 Participation_in_Infrastructures.txt
-- 0.113 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.112 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.111 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.110 Resolution_Theory.txt
-- 0.109 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.108 Machine_Pedagogies.txt

Machine_Listening.txt
-- 0.119 Machine_Pedagogies.txt
-- 0.115 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.116 Testing_Texting_South_a_Political_Fiction.txt
-- 0.113 An_Ethnography_of_Error.txt
-- 0.114 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.111 Relearn_to_Read_Speed_Readers.txt
-- 0.107 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.107 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.109 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.108 From_Page_Rank_to_Rankbrain.txt
-- 0.108 Participation_in_Infrastructures.txt
-- 0.102 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.102 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.101 Resolution_Theory.txt
-- 0.100 Machine_Pedagogies.txt

Machine_Pedagogies.txt
-- 0.119 Machine_Listening.txt
-- 0.118 An_Ethnography_of_Error.txt
-- 0.117 From_Page_Rank_to_Rankbrain.txt
-- 0.110 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.111 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.107 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.106 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.103 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.102 Testing_Texting_South_a_Political_Fiction.txt
-- 0.101 Relearn_to_Read_Speed_Readers.txt
-- 0.100 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.099 Resolution_Theory.txt
-- 0.098 Machine_Listening.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
-- 0.186 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.184 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.180 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.172 Testing_Texting_South_a_Political_Fiction.txt
-- 0.172 Relearn_to_Read_Speed_Readers.txt
-- 0.170 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.165 An_Ethnography_of_Error.txt
-- 0.163 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.161 Resolution_Theory.txt
-- 0.158 Machine_Listening.txt
-- 0.157 From_Page_Rank_to_Rankbrain.txt
-- 0.155 Machine_Pedagogies.txt

Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.138 From_Page_Rank_to_Rankbrain.txt
-- 0.133 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.127 Relearn_to_Read_Speed_Readers.txt
-- 0.126 Testing_Texting_South_a_Political_Fiction.txt
-- 0.122 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.121 Machine_Listening.txt
-- 0.118 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.117 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.116 Resolution_Theory.txt
-- 0.115 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.114 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.111 An_Ethnography_of_Error.txt

Relearn_to_Read_Speed_Readers.txt
-- 0.125 Testing_Texting_South_a_Political_Fiction.txt
-- 0.117 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.113 Machine_Listening.txt
-- 0.112 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.110 From_Page_Rank_to_Rankbrain.txt
-- 0.104 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.102 Participation_in_Infrastructures.txt
-- 0.102 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.100 An_Ethnography_of_Error.txt
-- 0.097 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.096 Machine_Pedagogies.txt
-- 0.095 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.092 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.091 Resolution_Theory.txt

Resolution_Theory.txt
-- 0.171 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.169 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.167 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.163 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.161 Participation_in_Infrastructures.txt
-- 0.159 Machine_Listening.txt
-- 0.147 Elusive_Borders_Aesthetic_Perpectualization_The_Space_Time_of_Metadata.txt
-- 0.145 An_Ethnography_of_Error.txt
-- 0.144 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.143 Relearn_to_Read_Speed_Readers.txt
-- 0.140 From_Page_Rank_to_Rankbrain.txt
-- 0.138 Testing_Texting_South_a_Political_Fiction.txt
-- 0.136 Machine_Pedagogies.txt
-- 0.134 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt

Testing_Texting_South_a_Political_Fiction.txt
-- 0.134 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.132 Relearn_to_Read_Speed_Readers.txt
-- 0.126 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.123 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.122 Machine_Listening.txt
-- 0.121 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.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_Runk_to_Rambrain.txt
-- 0.866 An_Ethnography_of_Error.txt
-- 0.861 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.855 Elusive_Borders_Aesthetic_Perspectualization_The_Space_Time_of_Metadata.txt
-- 0.849 Machine_Pedagogy.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
-- 0.893 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.892 Elusive_Borders_Aesthetic_Perspectualization_The_Space_Time_of_Metadata.txt
-- 0.869 Machine_Listening.txt
-- 0.868 Machine_Pedagogy.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_Testing_South_a_Political_Fiction.txt
-- 0.861 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.854 From_Page_Runk_to_Rambrain.txt
-- 0.852 Return_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
-- 0.134 Testing_Testing_South_a_Political_Fiction.txt
-- 0.115 Machine_Listening.txt
-- 0.892 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_Perspectualization_The_Space_Time_of_Metadata.txt
-- 0.884 Participation_in_Infrastructures.txt
-- 0.882 Return_to_Read_Speed_Readers.txt
-- 0.871 An_Ethnography_of_Error.txt
-- 0.871 Machine_Pedagogy.txt
-- 0.871 Resolution_Theory.txt
-- 0.866 Unmaking_Screens_a_Genealogy_of_the_Mineral_Vision.txt
-- 0.848 From_Page_Runk_to_Rambrain.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_Perspectualization_The_Space_Time_of_Metadata.txt
-- 0.893 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.891 From_Page_Runk_to_Rambrain.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_Testing_South_a_Political_Fiction.txt
-- 0.878 Machine_Listening.txt
-- 0.866 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.863 Return_to_Read_Speed_Readers.txt
-- 0.859 An_Ethnography_of_Error.txt
-- 0.859 Machine_Pedagogy.txt
-- 0.853 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
-- 0.893 Testing_Testing_South_a_Political_Fiction.txt
-- 0.882 The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.877 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.876 Machine_Listening.txt
-- 0.875 Elusive_Borders_Aesthetic_Perspectualization_The_Space_Time_of_Metadata.txt
-- 0.874 Return_to_Read_Speed_Readers.txt
-- 0.874 An_Ethnography_of_Error.txt
-- 0.873 Machine_Pedagogy.txt
-- 0.870 Computing_War_Narratives_The_Hamlet_Evaluation_System_in_Vietnam.txt
-- 0.866 The_Signification_Communication_Question_Some_Initial_Remarks.txt
-- 0.863 Resolution_Theory.txt
-- 0.861 The_Cultural_Politics_of_Information_and_of_Debt.txt
-- 0.854 From_Page_Runk_to_Rambrain.txt
```