abelardo\_g\_fournier\_2 abelardo\_g\_fournier\_2

## **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-<u>learn.org/stable/modules/feature\_extraction.html#text-feature-extraction)</u>

```
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)

```
CALCLATES SMILARITY VALUE STREEN DOCUMENTS

A. Elbourgarby, of farme tat

- 0.085 Mechine, listening, ste

- 0.085 Mechine, listening, ste

- 0.075 Mechine

- 0.075 
                               Computing_Mar_Narratives_The_Hamlet_Evaluation_System_in_Vietman.txt -- 0.081 Testing_Texting_South_a_Political_Fiction.txt
                                                  - 0.034 The_Cultural_Politics_of_Information_and_of_Debt.txt
- 0.033 The_Signification_Communication_Question_Same_Initial_Remarks.txt
- 0.024 Resolution_Theory.txt
                         - 0.028 Resolution, Theory, txt

| Clinicia, Border, Ashattic, Chryschauliantion, The Spoor, Time, of Metadate.txt
- 0.108 Perficiention, it, Infrastructures.txt
- 0.108 Perficiention, it, Infrastructures.txt
- 0.108 The Spring Meternt, that say, Ross, and, Love, Little, Americkial Menns.txt
- 0.108 The Schied Meternt, that say, Ross, and Love, Little, Americkial Menns.txt
- 0.102 The Colleged Meternt, Destructure, and Debt.txt
- 0.102 The Colleged Meternt Colleged C
                                                  - 0.091 The_stopic_Network_that_me_romo_ama_tow.txt
- 0.007 Maching_Modoppies.txt
- 0.000 Relation_to_Romd_Speed_Romders.txt
- 0.000 Relation_to_Romd_Speed_Romders.txt
- 0.000 Relation_title_South_political_Fiction.txt
- 0.007 Maching_title_South_political_Fiction.txt
- 0.007 Maching_title_South_political_Firet.txt
- 0.007 Goognating_Mor_Neurorities_The_Namelet_Folloation_System_in_Vietnon.txt
- 0.005 Computing_Mor_Neurorities_The_Indication_The_Speed_Time_of_Methodoto_txt

    0.857 Participation_in_Infrastructures.txt
    0.855 The_Cultural_Politics_of_Information_and_of_Debt.txt
    0.855 Unbuisting_Screens__Genealogy_of_the_Mineral_Vision.txt
    0.854 Unbuisting_Screens__Genealogy_of_the_Mineral_Vision.txt
    0.848 The_Signification_Communication_Question_Some_Initial_Remarks.txt
    0.848 Resolution_Theory_txt.
- 0.000 Resolution, Theory.txt

Rectine, Listening, tat

- 0.115 The, Signification, Communication, Question, Some_Initial_Remarks.txt

- 0.105 The, Signification, Communication, Question, Some_Initial_Remarks.txt

- 0.000 Resting_Texting_Loutus_Buth_u_Political_Fiction.txt

- 0.000 Recting_Texting_Loutus_Buth_u_Political_Fiction.txt

- 0.000 Recting_Texting_Loutus_Buth_u_Political_Fiction.txt

- 0.000 Recting_Texting_Loutus_Buth_u_Political_Fiction.txt

- 0.000 The Configuressum_Loutus_Buth_u_Political_Fiction.txt

- 0.000 Computing_Buth_u_Political_Fiction.txt

- 0.000 Computing_Buth_u_Political_Fiction.txt

- 0.000 Computing_Buth_u_Political_Fiction.txt
-- 0.80% Computing Nor-Norrotives, The Jamlet_Ivaluation_System_in, Vietnam.txt
Machine_Pologysis_txt
-- 0.135 Moction_Littering.txt
-- 0.808 Act Interpolys_of_Fore.txt
-- 0.808 Polity Terms_Polity_of_Fore.txt
-- 0.809 Polity_of_Fore_Polity_of_Fore.txt
-- 0.809 Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore_Polity_of_Fore
                                            - 0.188 branking. Screens.a, Genealogs.of.the.Blancal_Vision.txt
- 0.197 The.Schople, Metwork.thut.g. Nov., and_Love.tts
- 0.197 Elisave_Blancal_Vision.txt
- 0.196 Elisave_Blancal_Vision.txt
- 0.196 Elisave_Blancal_Vision.txt
- 0.196 Elisave_Blancal_Vision.txt
- 0.197 Elisave_Blancal_Vision.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-Amarchival-Means.txt
- 0.138 From.Page.Rank.tp. Amabbenin.txt
- 0.138 From.Page.Rank.tp. Amabbenin.txt
- 0.097 Relowers.Dead.Speed.Rank.dead.strian.The_Spore_Time_of_Metadata.txt
- 0.097 Relowers.Dead.Speed.Rank.dead.txtion.txt
- 0.097 Relowers.Dead.Speed.Rank.dead.txtion.txt
- 0.097 Relowers.Dead.Speed.Rank.dead.txtion.Speed.Thitiol.Jenewiks.txt
- 0.007 This_Spirit (rotion, Communication, Operation, Speed.Thitiol.Jenewiks, txt
- 0.008 Relowers.Dead.txtion.ty.Dead.txtion.txti
                         - 0.000 Profit-Composition. In, Infrastructures. Ltt.

- 0.000 Mochine-Pedagogies. Ltd.

- 0.007 Computing, Barry Pedagogies. Ltd.

- 0.007 Computing, Barry Pedagogies. Ltd.

- 0.007 Dissolution, Exerces. _ (Geneallogy of the Miseral _ Vision. Ltt.

- 0.000 Resolution. Theory _ Ltd.

- 0.000 Resolution. Theory _ Ltd.

- 0.000 Resolution. Theory _ Ltd.

- 0.000 The _ Ltd.

- 0.000 The
Belamin, B. Bedigmed, Broders, its

-0.15 Testing, Testing, South, p. Political, Fittion. txt

-0.15 Testing, Testing, South, p. Political, Fittion. txt

-0.687 Testime-Recognition-cross-foodies-and-Machines-by-Anarchival-Means, txt

-0.682 The Spirit Cettinn, Communication, Question_Some_Initial_Remarks, txt

-0.682 The Spirit Cettinn, Communication, Question_Some_Initial_Remarks, txt

-0.682 The Spirit Cettinn, Communication, Question_Some_Initial_Remarks, txt

-0.683 The Stupid_Members, Lendage_Initial_Remarks, Vision. txt

-0.687 Descripting initial_Information_Initial_Remarks

-0.682 The Stupid_Members, Lendage_Initial_Remarks

-0.682 Action_Initial_Remarks

-0.683 To Computing Nor. Numeratives, The Member_Loulustion_System_Initial_Remarks

-0.683 Members, Produpping Lendage_Initial_Remarks

-0.684 Members, Produpping Lendage_Initial_Remarks

-0.684 Members, Produpping Lendage_Initial_Remarks

-0.685 Members, Produpping_Initial_Remarks

-0.685 Members, Produpping_Initial_Rem
                               - 0.005 Participation, In, Infrastructures.txt
- 0.009 Mochine, intering.txt
- 0.009 Mochine, intering.txt
- 0.009 Mochine, intering.txt
- 0.000 Mochine, in
The_Stupid_Network_that_we_Know_and_Love.txt
-- 0.107 Participation_in_Infrastructures.txt
```

\_Love.txt .lization\_The\_Space\_Time\_of\_Metadata.txt

0.859 McChine\_Pedagogies.txt 0.859 Computing\_Wor\_Narratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt 0.844 Resolution\_Theory.txt

0.066 The\_Signification\_(ammunication\_Question\_Same\_Initial\_f
 0.063 Resolution\_Theory.txt
 0.061 The\_(ultural\_Politics\_of\_Information\_and\_of\_Debt.txt
 0.054 From\_Page\_Rank\_to\_Rankbrain.txt

CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS

## 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