abelardo\_g\_fournier

# 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-featureextraction (Proximus NV » Level 3 Communications, Inc. » GitHub, Inc.)

The calculations of the # word count of papers is based on work-in-progress versions of articles produced for the workshop. See https://machineresearch.wordpress.com/

### 

```
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)

 ${\tt Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt}$ -- 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)

#### CALCULATED SIMILARITY VALUE BETWEEN DOCUMENTS

An\_Ethnography\_of\_Error.txt

- 0.88 Mochine\_lespogles\_txt

- 0.88 Mochine\_lespogles\_txt

- 0.87 Mochine\_listering\_txt

- 0.867 From\_Poge\_Rank\_to\_Rankbroin.txt

- 0.867 From\_Poge\_Rank\_to\_Rankbroin.txt

- 0.868 Fasting\_Texting\_South\_APOlitical\_liction.txt

- 0.868 Fasting\_Texting\_South\_APOlitical\_liction.txt

- 0.869 Fasting\_Texting\_South\_APOlitical\_liction.txt

- 0.869 Fasting\_Texting\_South\_APOlitical\_liction.txt

- 0.869 Fasting\_Texting\_South\_APOlitical\_liction.txt

- 0.869 Taxing\_Texting\_txt

- 0.869 Taxing\_Texting\_South\_APOlitical\_liction\_Texting\_txt

- 0.874 Filest\_Nonders\_Aresthict\_Perpectual\_liction\_Texting\_txt

- 0.868 Camputing\_Non\_Narroritives\_The\_Hamlet\_Evaluation\_System\_in\_Vietnon.txt

- 0.868 Taxing\_Uniton\_Theory\_txt

-

Computing\_Nor\_Norrotives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt
-- 0.881 Testing\_Texting\_South\_a\_Political\_Fiction.txt
-- 0.878 Pattern-Recognition\_cross-Bodies\_and-Machines-by-Anarchival-Means.txt
-- 0.8780 Puttern-Recognition\_cross-Bodies\_and-Machines-by-Anarchival-Means.txt
-- 0.8780 Participation\_in\_Linfrastructures.txt
-- 0.8780 Participation\_in\_Linfrastructures.txt

From Page, Bank, to, Bankbrain.txt

- 0.138 Pattern-Recognition-across-Badies-and-Machines-by-Anarchival-Means.txt

- 0.091 The Stupid Alebanch tybol.new.fxom.and\_Love.txt

- 0.097 Machine Pedagogies.txt

- 0.098 Pachine Pedagogies.txt

- 0.098 Pachine Depad Speed Readers.txt

- 0.098 Pachine Losening South\_L\_Political\_Fiction.txt

- 0.096 Machine\_Listening.txt

- 0.066 Machine\_Listening.txt

- 0.066

- 0.067 An\_Ethiography\_of\_Error\_tst
- 0.065 Computing\_Bur Pairortives\_Ne\_Maint\_Evaluation\_System\_in\_Vietmon\_txt
- 0.066 Elusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metodata\_txt
- 0.069 Tarticipation\_in\_Informaticutures\_txt
- 0.057 Participation\_in\_Endocutures\_txt
- 0.054 Phracking\_Screens\_o\_demongly\_of\_the\_Minrol\_Vision\_txt
- 0.084 thracking\_Screens\_o\_demongly\_of\_the\_Minrol\_Vision\_txt
- 0.084 thracking\_Screens\_o\_demongly\_of\_the\_Minrol\_Vision\_txt
- 0.086 Residuation\_Theory\_txt

- w.ww Mesoution\_Imcory.txt

- 0.135 Methins\_hologogies.txt
- 0.135 Methins\_hologogies.txt
- 0.135 Methins\_hologogies.txt
- 0.135 Methins\_hologogies.txt
- 0.085 An\_Elbrins\_hologogies.txt
- 0.085 An\_Elbrins\_hologogies.txt
- 0.085 An\_Elbrins\_firetion\_comming\_tion\_location\_firetion.txt
- 0.085 An\_Elbring\_robe\_in\_Elbrins\_troc\_txt
- 0.085 Mething\_robe\_in\_Elbrins\_troc\_txt
- 0.085 Mething\_robe\_in\_to\_location\_location\_location\_location\_txt
- 0.087 Damaking\_Screens\_0\_demology\_of\_the\_Minerol\_Vision.txt
- 0.087 To\_no\_tion\_in\_Elbrins\_troc\_troc\_txt
- 0.087 Methins\_troc\_troc\_txt
- 0.085 From\_Toge\_Bonk\_to\_Robe\_for\_txt
- 0.085 From\_Toge\_Bonk\_to\_Robe\_for\_txt
- 0.085 Prot\_troc\_troc\_txt
- 0.085 Methins\_troc\_troc\_txt
- 0.085 Methins\_troc\_troc\_txt
- 0.085 Methins\_troc\_txt
- 0.085 Methins\_troc\_txt
- 0.085 Methins\_txt

-- 0.836 Computing\_Mor\_Marratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt
Mochine\_Pedagogies\_txt
-- 0.8119 Machine\_Listening\_txt
-- 0.883 An\_Ethnography\_of\_Error\_txt
-- 0.883 An\_Ethnography\_of\_Error\_txt
-- 0.883 From\_Dage\_Mank\_tc\_Mansharin.txt
-- 0.883 From\_Dage\_Mank\_tc\_Mansharin.txt
-- 0.887 From\_Dage\_Mank\_tc\_Mansharin.txt
-- 0.897 Ime\_Signification\_Communication\_Outsion\_Some\_Initial\_Remarks\_txt
-- 0.897 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks\_txt
-- 0.863 The\_Cultural\_Politics\_of\_Information\_ond\_of\_Debt\_txt
-- 0.863 The\_Cultural\_Politics\_of\_Information\_ond\_of\_Debt\_txt
-- 0.863 The\_Cultural\_Politics\_of\_Information\_Inversety
-- 0.865 Resignification\_South\_Debt\_tics\_of\_Information\_Inversety
-- 0.865 Terric\_Detain\_On\_Deltics\_of\_Initial\_Cition\_txt
-- 0.865 Terric\_Detain\_On\_Deltics\_of\_Initial\_Cition\_txt
-- 0.865 Terric\_Detain\_on\_In\_Information\_Unitial\_Cition\_System\_in\_Vietman.txt
-- 0.865 Resolution\_Theory\_txt

Potern-Recognition-cross-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.138 From.Page.Rank.to.Bankbrain.txt

- 0.138 From.Page.Rank.to.Bankbrain.txt

- 0.108 Situsive.Banders\_Assetheti.Perpectualization.The\_Space\_Time\_of\_Metadata.txt

- 0.097 Relearn.to.Read.Speed.Readers.txt

- 0.096 Testing\_Texting\_South\_Political\_Fittion.txt

- 0.096 Testing\_Texting\_South\_Political\_Fittion.txt

- 0.084 Machine\_Listening.txt

- 0.088 Machine\_Listening.txt

- 0.088 Machine\_Madagagies.tx

- 0.088 Machine\_Madagagies.txs

- 0.088 Machine\_Madagagies.txs. The\_Mealet\_Follution\_System\_in\_Vietnon.txt

- 0.098 Machine\_Madagagies.tx

- 0.098 Machine\_Office\_Time\_Madagagies.tx

- 0.098 M

-- 0.061 An\_Ethnography\_of\_Error.txt

Relearn\_to\_Read\_Speed\_Readers.txt
-- 0.125 Feating\_Texting\_South\_a\_Diltical\_Fiction\_txt
-- 0.097 Pattern-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
-- 0.083 Machine\_Listering\_txt
-- 0.0827 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt
-- 0.0827 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt
-- 0.0837 For\_DapE\_Bank\_to\_Robarborin.txt
-- 0.0927 Participation\_In\_Infrastructures.txt
-- 0.0927 Participation\_In\_Infrastructures.txt
-- 0.0927 Participation\_In\_Infrastructures.txt
-- 0.0937 The\_Signification\_Initial\_Remarks.txt
-- 0.0937 Tomputing\_Mar\_Instructures.txt
-- 0.0937 Tomputing\_Mar\_Instructures.txt
-- 0.0937 Tomputing\_Mar\_Instructures.txt
-- 0.0937 Tomputing\_Mar\_Instructures.txt
-- 0.0938 The\_Cithography.of\_Error.txt
-- 0.0938 The\_

Testing\_Texting\_South\_a\_Political\_fiction.txt

-- 0.134 The\_Signification\_Communication\_Question\_Some\_Initial\_Remarks.txt

-- 0.125 Relearn\_to\_Read\_Speed\_Readers.txt

-- 0.969 Pattern-Recognition\_coros=Bodies=and-Machines-by-Anarchival-Means.txt

-- 0.999 Jumaking\_Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt

-- 0.989 Machine\_Listening\_txt

-- 0.889 Machine\_Listening\_txt

-- 0.898 Machine\_Listening\_txt

The Loutural Doilties of Information.and. of Debt.txt

- 0.164 The Signification.Communication.Question.Some\_Initial\_Remorks.txt

- 0.164 The Signification.Communication.Question.Some\_Initial\_Remorks.txt

- 0.093 The Subuid\_Aktwork\_thot.ee.Xnow.and\_iover.tst

- 0.095 Zelusive\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadata.txt

- 0.068 Mochine\_Istatening.txt

- 0.067 Resolution\_Theory.txt

- 0.067 Resolution\_Theory.txt

- 0.067 Resolution\_Theory.txt

- 0.063 Participation\_Initial\_Transructures\_txt

- 0.063 Testing\_Stating\_South\_a\_Political\_Transructures\_txt

- 0.063 Testing\_Stating\_South\_a\_Political\_Transructures\_txt

- 0.063 Testing\_Stating\_South\_a\_Political\_Transructures\_txt

- 0.063 Testing\_Stating\_Statens\_Centens\_Gov\_off\_the Mineral\_Vision.txt

- 0.063 Descript\_Stating\_Statens\_Centens\_Gov\_off\_the Mineral\_Vision.txt

- 0.063 Descript\_Stating\_Statens\_Centens\_Gov\_off\_the Mineral\_Vision.txt

- 0.063 Descript\_Stating\_Stating\_Stating\_Testing\_Centens\_Gov\_off\_the Mineral\_Vision.txt

- 0.063 All Angle Testing\_Stating\_Testing\_Sta

The Signification Communication Question Some Initial Remarks.txt

- 0.164 The Cultural Politics of Information and of Debt.txt

- 0.134 Testing Texting Souths, Political Fittion:txt

- 0.135 Testing Texting Souths, Political Fittion:txt

- 0.135 Machine\_Listening.txt

- 0.092 Pattern-Recognition-ocross-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.092 Pattern-Recognition-ocross-Bodies-and-Machines-by-Anarchival-Means.txt

- 0.095 The Stupid Network\_thot\_me\_Knom\_and\_Love.txt

- 0.095 The Stupid Network\_thot\_me\_Knom\_and\_Love.txt

8.082 Umaking\_Screens\_0\_Gemeology\_of\_the\_Minerol\_Vision.txt
0.078 Testing\_Festing\_South\_o\_Political\_Fiction\_Vision.txt
0.070 Nochine\_Listening\_txt
10.070 Nochine\_Listening\_txt
0.065 Pattern\_Recognition\_across-Bodies\_and-Machines-by-Anarchival\_Means\_txt
10.063 Relearn\_to\_Read\_Speed\_Readers\_txt

0.059 Machine\_Pedagogies.txt
- 0.059 Machine\_Pedagogies.txt
- 0.059 Computing\_Nar\_Marratives\_The\_Hamlet\_Evaluation\_System\_in\_Vietman.txt
- 0.044 Resolution\_Theory.txt

Screens\_a\_Genealogy\_of\_the\_Mineral\_Vision.txt
Participation\_in\_Infrastructures.txt

188 wortstipation in infrastructures.txt
907 Esting Texting South, 2011 listical filest.txt
1877 Potterm-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
1877 Fotterm-Recognition-across-Bodies-and-Machines-by-Anarchival-Means.txt
1875 Tlusture\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadota.txt
1875 Tlusture\_Borders\_Aesthetic\_Perpectualization\_The\_Space\_Time\_of\_Metadota.txt
1874 Relearn.to\_Bod\_Speed\_Readors.txt

0.866 The\_Signification\_Communication\_Question\_Some\_Initial 0.863 Resolution\_Theory.txt 0.861 The\_Cultural\_Politics\_of\_Information\_and\_of\_Debt.txt 0.854 From\_Page\_Rank\_to\_Rankbrain.txt

## abelardo\_g\_fournier

## 

#### 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