GORDON AND BETTY MOORE FOUNDATION

GitHub

CITEMAP

Mayank Kedia, Kris Kooi, Chrisantha Perera, Flip Tanedo

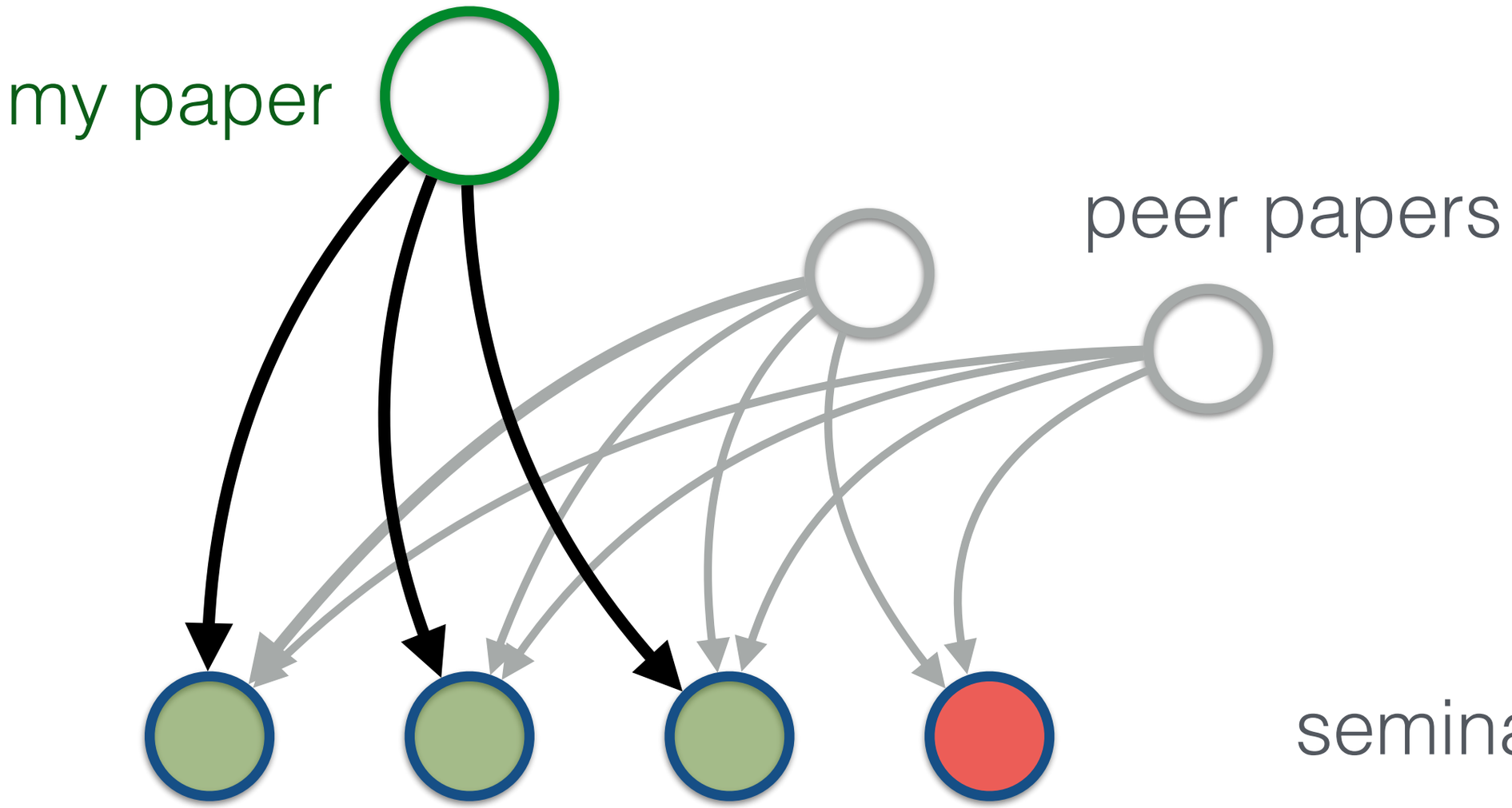
git.io/vWatB

With thanks to: *mitar*, Tadej Novak, *Steven !Ragnarök*, *ivan*, *dreww*, Julia Bossmann, Julia Graber, Ryan Anderson, & everyone who chatted with us!

SCIENCE HACK DAY

San Francisco 2015

The Problem


my paper
peer papers
seminal papers
papers that I cited
forgot to cite

The Data Set

Open access literature database for high energy physics

Clean: Remove empty data, incomplete data (old papers)
Make tractable: only take papers after 2000

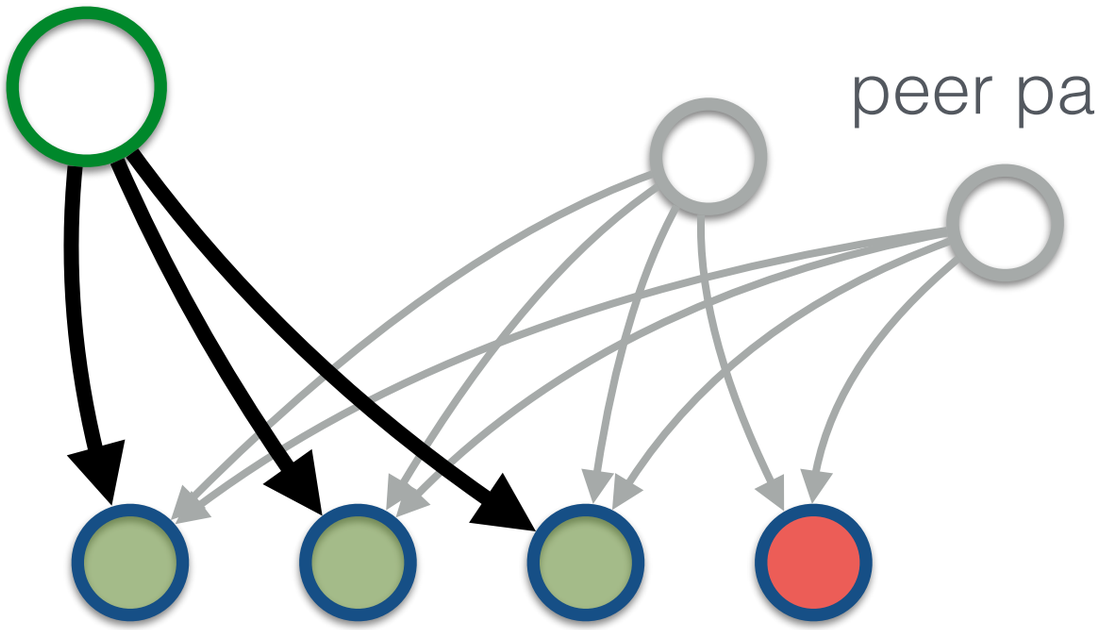
~ 500,000 entries


```
{
  "abstract": {
    "summary": "We complete the effective potential calculation of the two-loop, top/bottom Yukawa
corrections to the Higgs boson masses in the Minimal Supersymmetric Standard Model, by computing the O(at^2
+ at*ab + ab^2) contributions for arbitrary values of the bottom Yukawa coupling. We also compute the
corrections to the minimization conditions of the effective potential at the same perturbative order. Our
results extend the existing O(at^2) calculation, and are relevant in regions of the parameter space
corresponding to tan(beta) >> 1. We extend to the Yukawa corrections a convenient renormalization scheme,
previously proposed for the O(ab*as) corrections, that avoids unphysically large threshold effects
associated with the bottom mass and absorbs the bulk of the corrections into the one-loop expression. For
large values of tan(beta), the new contributions can account for a variation of several GeV in the lightest
Higgs boson mass."
  },
  "recid": 618609,
  "title": {
    "title": "On the two loop Yukawa corrections to the MSSM Higgs boson masses at large tan beta"
  }
}
```

Collaborative Filtering

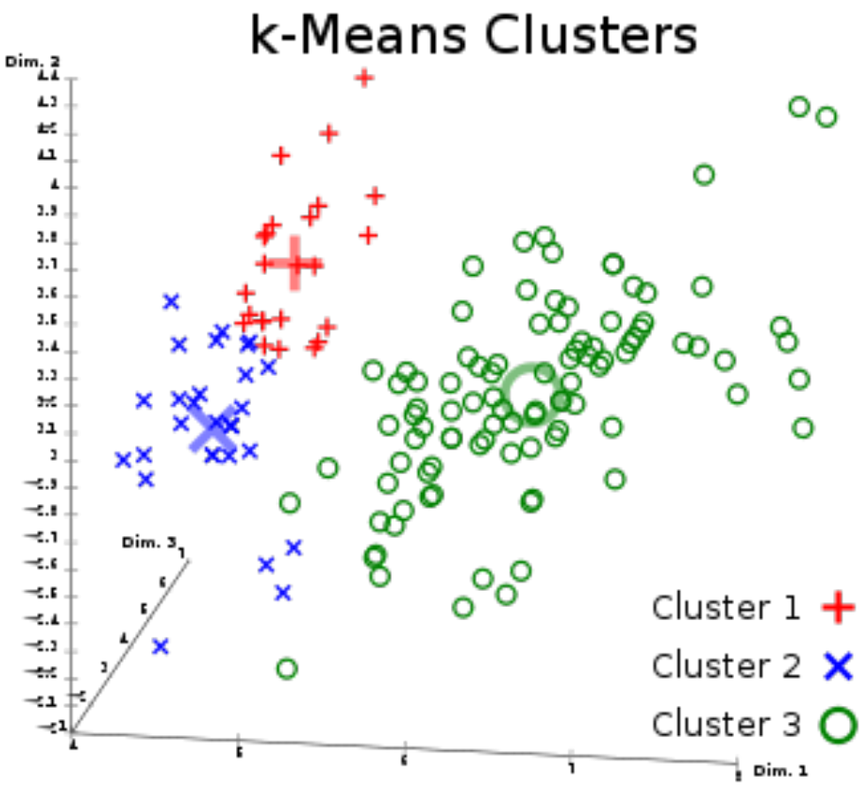
Using Peer Papers

RULES-BASED



Bag of Words

K-MEANS CLUSTERING



Wikimedia, Chire, *k-means*

Example

The First VLT FORS1 spectra of Lyman-break candidates in the HDF-S and AXAF deep field
S. Cristiani *et al.*. Apr 2000. 5 pp.
Published in **Astron.Astrophys. 359 (2000) 489**
ESO-1378
e-Print: astro-ph/0004213 | PDF

Statistical properties of ultraluminous iras galaxies from an hst imaging survey
J. Cui, X.Y. Xia, Z.G. Deng, S. Mao, Z.L. Zou. Apr 2001. 38 pp.
Published in **Astron.J. 122 (2001) 63**
DOI: 10.1086/321127
e-Print: astro-ph/0104296 | PDF

Local Lyman Break Galaxy Analogs: The Impact of Massive Star-forming Clumps on the Interstellar Medium and the Global Structure
R.A. Overzier *et al.*. Oct 2009. 22 pp.
Published in **Astrophys.J. 706 (2009) 203-222**
DOI: 10.1088/0004-637X/706/1/203
e-Print: arXiv:0910.1352 [astro-ph.CO] | PDF

Cosmological Galaxy Formation Simulations Using SPH
G. Stinson, J. Bailin, H. Couchman, J. Wadsley, S. Shen, C. Brook, T. Quinn. Apr 2010. 16 pp.
Published in **Mon.Not.Roy.Astron.Soc. 408 (2010) 812**
DOI: 10.1111/j.1365-2966.2010.17187.x
e-Print: arXiv:1004.0675 [astro-ph.CO] | PDF

Possible Directions

In Vivo testing with academics

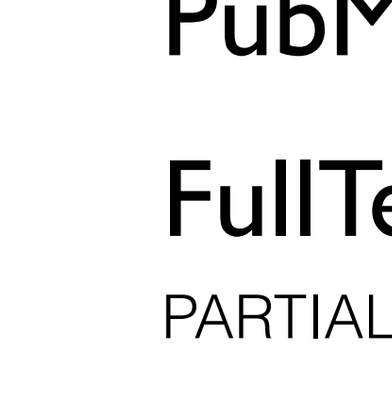
PubMed/CiteSeer^x dataset

FullText Analysis

PARTIAL DATASET: bit.ly/1RvuK4F

Heroku web service

Run on Spark



... and continue to spread & hack

SCIENCE HACK DAY

thanks everyone!