

# DATASCI W261: Machine Learning at Scale

## Group 3, Homework 13

**Konniam Chan** (konniam.chan@berkeley.edu)

**Nick Hamlin** (nickhamlin@gmail.com)

**Tigi Thomas** (tgthomas@berkeley.edu)

**Angela Gunn** (angela@egunn.com)

**Jaime Villalpando** (jaimegvl@hotmail.com)

**Safyre Anderson** (think2twice.saf@gmail.com)

**Karthik Chepudira** (kchepudira@berkeley.edu)

**Jing Xu** (jing.xu@ischool.berkeley.edu)

**Yi Jin** (yjin@ischool.berkeley.edu)

**Vamsi Sakhamuri** (vamsi@ischool.berkeley.edu)

**Howard Wen** (howard\_wen@ischool.berkeley.edu)

**Prabhakar Gundugola** (prabhakar@ischool.berkeley.edu)

Time of Submission: 9:00 AM EST/6:00 AM PST, Friday, April 29, 2016

W261-3, Spring 2016

## Useful References and Notebook Setup:

- **Original Assignment Instructions** (<https://www.dropbox.com/s/gsti4plbst7ena3/MIDS-MLS-HW-13.txt?dl=0>)

## Subnotebooks

We've taken a divide and conquer approach on this assignment, so the different portions of the assignment are available via the linked notebooks below:

- **13.1-13.3: Pagerank Local Content (in Jupyter)** ([https://github.com/KonniamChan/machine-learning-at-scale/blob/master/Week\\_13/MIDS-W261-2016-HWK-Week13-Group3\\_PageRank.ipynb](https://github.com/KonniamChan/machine-learning-at-scale/blob/master/Week_13/MIDS-W261-2016-HWK-Week13-Group3_PageRank.ipynb))
- **13.1-13.3: Pagerank EMR Content (in Zeppelin)** (<https://www.zeppelinhub.com/viewer/notebooks/aHR0cHM6Ly9yYXcuZ2l0aHVidXNlcmNvbnRlbnQu>)
- **13.4-13.5 Criteo (in Jupyter)** ([https://github.com/nickhamlin/mids\\_261\\_homework/blob/master/HW13/Group3\\_HW13\\_Criteo.ipynb](https://github.com/nickhamlin/mids_261_homework/blob/master/HW13/Group3_HW13_Criteo.ipynb))
- **13.6 (OPTIONAL) HHH Problem (in Jupyter)** ([http://nbviewer.jupyter.org/urls/dl.dropbox.com/s/hyliycz9dv8mso7/MIDS\\_W261\\_HW13\\_Angela\\_Gun](http://nbviewer.jupyter.org/urls/dl.dropbox.com/s/hyliycz9dv8mso7/MIDS_W261_HW13_Angela_Gun))

In [ ]: