

Applied Machine Learning

W207

datascience@berkeley

Week 1 Agenda

- How this class works: orga, syllabus, etc
- Introductions - Get to know each other
- Example notebook
- ‘Unreasonable Effectiveness’ paper

How this class works: orga, syllabus, etc

datascience@berkeley

A word on COVID-19

datascience@berkeley

Communication

- Email: dschib@berkeley.edu → for personal non-tech questions only
- Slack channel for this class: #w207_schioberg → use this for code/content questions, highest chance for quick help, or use
- The "big" channel: #w207
- Office hours: Tuesdays 8:05 PM PST or by appointment
 - You can go to other instructors' office hours. Should be shared in ISVC
- Announcements: Slack only! #w207_schioberg_announce - Double check your slack notification settings

Syllabus (in Github) - general approach

- One algorithm each week, will dig deeper in a few of them.
- Typical class outline (may vary based on topic):
 - Review async material
 - Walk through notebooks or small group work on notebooks (or both)
 - Dig deeper on the algo of the week, examples, questions
 - Discuss a paper/reading if time allows (will be announced beforehand)
- Find readings here:
<https://github.com/MIDS-W207/coursework/tree/master/Readings>
- Find Syllabus here:
https://github.com/MIDS-W207/coursework/blob/master/Schioberg/datasci-w207_syllabus.pdf

datascience@berkeley

I need to give you access to the github repo!!!
Fill out the survey pinned to slack

Syllabus (in github) - details

Week 1: Welcome!

Week 2: Nearest Neighbors

Week 3: Naive bayes, Spam Classification

Week 4: Decision Trees, Bagging, Boosting

Week 5: Linear Regression, Logistic
Regression.

Week 6: Gradient Descent, Regularization
(Deep Learning)

Week 7: Neural Networks (Deep Learning)

Week 8: Algorithm Comparison (Deep Learning)

Week 9: K-Means

Week 10: Gaussian Mixture Models

Week 11: PCA

Week 12: Graph Analysis

Week 13: Recommender Systems

Week 14: Class presentations (project 4)

I need to give you access to the github repo!!!
Fill out the survey pinned to slack

datascience@berkeley

Projects, Grades, Rules, Hints (1)

- **ALL TIMES (due dates, office hours, etc) IN THIS CLASS ARE PST**
- Grades are based on 4 projects. See syllabus (in github, folder Schioberg) for grading scheme
- 3 individual Python projects - work in groups, submit individually
- 1 group **kaggle competition** presented in the last lecture as a group. Details will follow
- Due dates: Sunday night (23:59 PST) after weeks 5, 9, 12 (exact dates to follow).
- Late submission = -10% on the grade! General rule for all 207 sections
- Advanced ML/programmer? It might still take you longer than expected :)

Projects, Grades, Rules, Hints (2)

- All projects available now in Github. You can start any of them any time. (I need to add you!)
- Project 1 available in GitHub classroom: <https://classroom.github.com/a/dCA1Y02o>
 - no access needed, just a github account
 - link also pinned to Slack channel
- Use Github whenever possible!
- Projects are graded by a TA/me by hand: I **read** and **run** everyone's **code**. Be ready to explain your setup to me so I can recreate it and see if your code actually runs.
- Upload your finished notebook to ISVC!
- ISVC does not allow re-uploads :(Contact me if you submitted too early.
- COMMENT as much as possible in your code! Explain your train of thoughts to me -> extra points even if code looks wonky and result is off

FAQ

- Can I get an extension? Yes, but with a **10% penalty**. Blanket rule in 207
- How do I turn in work? Upload your notebook to ISVC
- Best place for questions? Slack is best (general w207 and per section).
- Groups? YES PLEASE but submit individual solutions! In doubt you have to be able to explain your solution to me.
- Semester textbooks? ALL recommended readings are here:
<https://github.com/MIDS-W207/coursework/tree/master/Readings>
 - Want more? Two (NOT mandatory) recommendations
<http://www.deeplearningbook.org/> <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>

Your questions?

- Ask more questions

datascience@berkeley

"This doesn't run... HELP!"

How to ask for help

In case of code bugs!

1. Send code as a code snippet in slack. Please do NOT screenshot your notebook/shell/script!
2. Say what you were trying to do with some details.
3. Where are you trying to run this? Local Jupyter installation, colab, GCloud etc
4. Copy paste the whole error message into a code snippet in Slack

Why all this? I want to understand what you wanted to do, reproduce the error, and give you a helpful answer quickly

No screenshots! Really!

All other questions: simply explain where you are stuck! (if possible without screenshots)

datascience@berkeley

Introductions

[datascience@berkeley](mailto:datascience@berkeley.edu)

Introductions

- Tell us about yourself - your background, location, etc
- What (topic) are you most excited about in this class?
- Fun fact about yourself?

What's Artificial Intelligence?

datascience@berkeley

How to code in this class

Tutorial.ipynb is found in github

<https://github.com/MIDS-W207/coursework/blob/master>Notebooks/Week1>Notebooks/Tutorial.ipynb>