Effective Note-taking

* Bullet journaling
* Sketchnotes: draw in your notes
* Leave a page for index
* Make connections in your mind
* Stop highlighting or underlining the notes, it erases the context of points and makes difficult to form connections
* Re-read your notes to refresh your memory
* Avoid too many notebooks

File management

* Proper named and numbered folders
* Back up your data regularly
* Stop redundancy
* Find efficient data storage places such as server, cloud, drive etc.

Productivity

* Stop Multi-tasking
* Sprints of focused work-20 mins rule
* Accumulate small tasks (less creative) and do them all at one go, eg, checking emails etc
* Connect the things you study, find a relationship
* Don’t let others destroy your focus
* First plan, and then do that work with focus. Focus is the most important skill
* Prioritize your work: Don’t take new problems until previous ones are solved, first things first

Research Skills

* Understand the problem statement very well
* Select the right papers: based on credibility and relevance
* Summarize the papers in excel file(very important, else you will forget later). Use a research template
* Read all technical terms , jargons related to your field through abstracts and websites and create a dictionary in your mind(no need to know everything in detail)
* Divide the research in two parts: 1) Paper reading 2) General information online
* Set deadlines
* Read right papers: Journal, publication date, relevance , author affliation
* Read breadth first: broad idea by skimming through headlines, abstract introduction, can I apply this
* Read Depth second: Understand the method, assumptions, errors statistics, conclusions
* Understand the basic experimental set up of the problem
* Questions to ask while reading the paper:
  + What the main idea?
  + What are the assumptions?
  + Are the assumptions reasonable?
  + Does your problem follows the same assumption?
  + What are statistical test and error stats used to examine the method?
  + Does this follow the general experimental set up?
  + Is the method reducing the error significantly compared to other conventional methods?
  + Do the conclusions logically follow the observations?
  + What can be other explanations of the observations?
  + Is there any interesting observation left out by the author?
* Questions to ask after reading the paper:
  + Whats the new idea?: extract the innovation
  + What are the major points ? and summarize them
  + Can I use this method or any part of this method?
* Use other resouces (other than research papers): Abstracts, ppts, video, demo, code, pre-prints, blogs

Research methodology:

1. Understand the problem
2. Find 2 recent review paper and last 5 year papers
3. While reading paper: focus on method and result part(always ask ‘why?’)
   1. Ask above questions: main idea, assumptions, datasets, experimental setup, error stats, and improvement from conventional, ie. Problems solved,
   2. Summarize in 3 lines
   3. Important conclusions/observations
   4. Future work/scope of improvement
   5. Save the block diagram
   6. Think if any part is useful for u
4. After reading around 10 papers create a general block-diagram of the methods and a list of keywords
5. Read more papers and Find gaps: problems not yet considered, data not yet explored, experiments not yet performed, select best paper to build upon
6. Create hypothesis (for specific situations)
7. Prepare data and experiment setup
8. Do experiments and Prove hypothesis
9. Analyse the results- effectiveness, scope of the results, how reliable are they, are the assumptions u consider justified
10. Compare the result with a conventional method: decide which conventional method to compare with and what error metrics
11. Generalize the results
12. Future work

Note: Keep a log of all experiments and conclusions

Good programming practices:

1. Proper formatting
2. Uniform naming convention
3. Modular programming
4. Proper commenting
5. Portability: Not hard-coded
6. Proper arrangement of working folders and data

Image Processing and CV

1. Scikit-image tutorial
2. Python crash course
3. Revise image processing(from BKM and Gonzalez book) and select algorithms to code
4. Learn Computer vision (from video tuts and book) and select algorithms to code

Platform to use: Python, openCV, scikit-image, scikit-learn

How to create data science port-folio:

Some ways to push your boundaries are:

* Try working with a larger dataset than you’re comfortable with
* Start a project that requires knowledge you don’t have
* Try making your project run faster
* See if you can teach what you did in a project to someone else

**Types of projects**

DATA CLEANING

Data story telling

End to end

Explanatory post

* Ability to communicate
* Ability to collaborate with others
* Technical competence
* Ability to reason about data
* Motivation and ability to take initiative

Email management

* + Set up 3 email accounts: work, personal, bulk
  + Schedule a time for email check and keep your inbox zero
  + Do not check emails while on the go in lift, or line: reduces your ‘be in the moment’ activity, let yourself daydream
  + Keep work email only in the office and not in personal devices, else you are kept occupied in the office work
  + Reduce your time of email reading and responding
  + Unsubscibe for useless subscriptions

Productive breaks

* + Take a five minute to 10 minute break after doing 40 mins focused work
  + Things to do in a break:
    - Prepare snack
    - Give head/neck massage
    - Read an interesting article
    - Watch ted talk
    - Make coffee
    - Clean space around
    - Play mobile games like lumosity
    - Go out for a walk
    - Call a family member
    - Listen to a podcast
    - Hangout with a friend
    - go out for lunch
    - book a massage
    - play a sport with friend
    - sit in a café and watch people

Filing a patent (advice by Aoki-san):

Think about a good technology, contributing to the solution for a certain real problem, before thinking about a good patent.

Basically good technology is a seed of good patent.

It is difficult to provide the clear definition of good technology, but in my opinion, if you find a new approach based on a scientific principle and/or findings from new viewpoints, rather than superficial/empirical improvement of an existing technique, it is a good technology. In most cases, some superficial/empirical improvements are *additionally* needed toward realizing the application using our technology, though.

Additionally, newer technology based on newer principle can basically produce more derivative patents. If the contribution of the technology to the actual business is significant, those patents get regarded as good patents.

Regards,

Hirofumi

Being able to make the most out of very little data is a key skill of a competent data scientist.

Goal of weekly report:

The important point is what purpose you are doing each activity for, and what you do next considering your goal/milestone.

* 1. Even if the idea is in thinking phase, write briefly about it and describe that you will brush up your idea through the discussion with members from now
  2. Finding issues is also regarded as progress
  3. For weekly report of team, write ‘truth’ about current status, no need to hide details
  4. generally for the technology development, there are almost no cases where the first idea solely produce ideal results. It is generally needed to improve the idea and/or introduce new ideas and sometimes to change the direction, toward achievement of good results enough to be profitable in actual businesses
  5. include information of both achievements and next direction in title as much as possible. Instead of writing general research theme as title, write specific progress in title