Methods 3: Multilevel Statistical Modeling and Machine Learning

Week 11: *Final evaluation and wrap-up of course*
December 7, 2021

by: Lau Møller Andersen

These slides are distributed according to the CC BY 4.0 licence:
https://creativecommons.org/licenses/by/4.0/

December 16th
(11-13)

To wear:
Sports clothes
Preferably no zippers
No jewellery
Please also have a valid Corona passport

Silkeborgvej 41E, baghuset 1.th, 8000 Aarhus C.

Sign-up list here:
https://cryptpad.fr/pad/#/2/pad/edit/NHEOwwpF2nH8vTGDNPuQZENx/

EXAM

main ▾ **github_methods_3** / **exam** /

Go to file | Add file ▾ | ···

ualsbombe exam assignments — e980366 4 minutes ago — **History**

..		
assignment_1.Rmd	exam assignments	4 minutes ago
assignment_1.pdf	exam assignments	4 minutes ago
assignment_2_1.Rmd	exam assignments	4 minutes ago
assignment_2_1.pdf	exam assignments	4 minutes ago
assignment_2_2.Rmd	exam assignments	4 minutes ago
assignment_2_2.pdf	exam assignments	4 minutes ago
assignment_3.Rmd	exam assignments	4 minutes ago
assignment_3.pdf	exam assignments	4 minutes ago
assignment_4.Rmd	exam assignments	4 minutes ago
assignment_4.pdf	exam assignments	4 minutes ago

IMPORTANT:

Make sure to submit a (knitted) *pdf (*or *html)* that shows the plots and the (printed) content of the variables.

No *Rmd* files!

Course titel	Exam form	Form of co-examination	Exam dates	Assessment deadline	Register / unregister for re-exam
BA 3. semester 2020					
Methods 3: Multilevel Statistical Modeling and Machine Learning	Portfolio	Internal Passed /failed	Submission: 13/12	10/1	13/1

Correction – assignment 3, exercise 1.1.iv and 1.2.ii

iv. Make an average over the repetition dimension using `np.mean` - use the `axis` argument. (The resulting array should have two dimensions ~~with time as the first and magnetic field as the second~~ with sensor as the first and time as the second)

ii. Now make four averages (As in Exercise 1.1.iii), one for each PAS rating, and plot the four time courses (one for each PAS rating) for the sensor found in Exercise ~~1.1.v~~ 1.1.vi

reffering to the wrong exercise

Correction – assignment 4

EXERCISE 2 - Use logistic regression with cross-validation to find the optimal number of principal components

1) We are going to run logistic regression with in-sample validation
   i. First, run standard logistic regression (no regularization) based on ~~$Z_{d\times k}$~~ and $Z_{n\times k}$ `y` (the target vector). Fit (`.fit`) 102 models based on: $k = [1, 2, ..., 101, 102]$ and $d = 102$. For each fit get the

2

classification accuracy, (`.score`), when applied to ~~$Z_{d\times k}$~~ and $Z_{n\times k}$ and y. This is an in-sample validation. Use the solver `newton-cg` if the default solver doesn't converge

AT AFLEVERE SOM STUDIEGRUPPE!!!!

Så vi kan godt skrive at "denne opgave har vi skrevet sammen"

Evaluation

Overall good

... but room for improvement

Response rate

Response rate

	Student
Responded	30
Invited	35
Response Ratio	85.71%

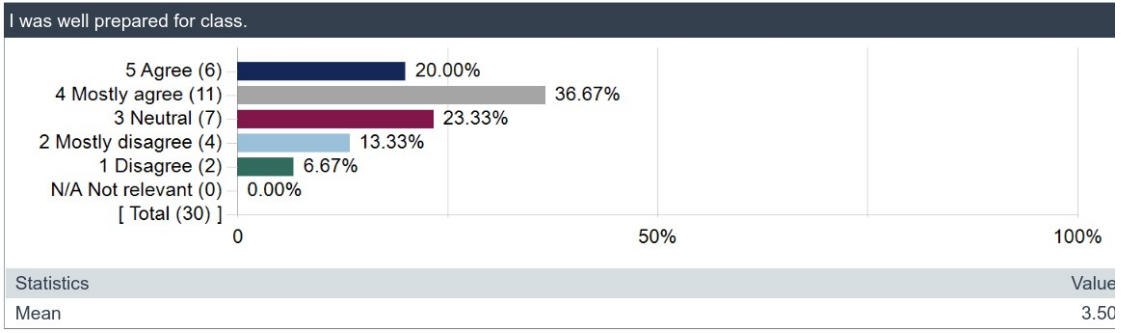
Summary

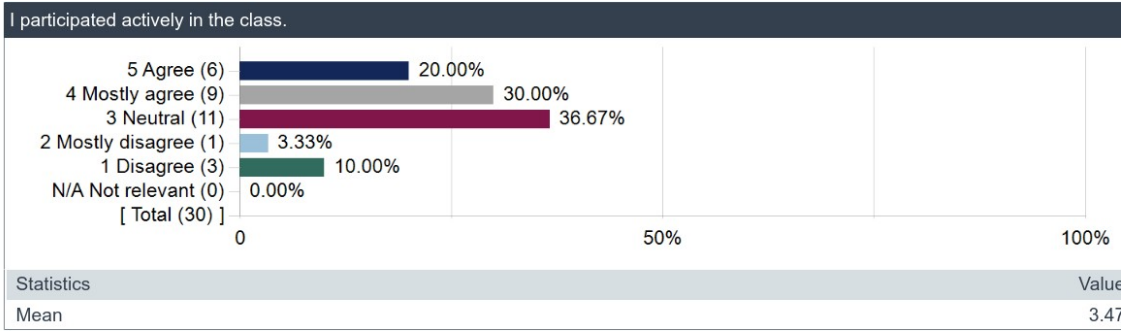
Course questions

Mean (likert scale 1-5)

	Mean
I rate the overall outcome of the course as:	3.33
In this course, the use of digital study and teaching activities (supported by eg Peergrade, Mentimeter, Padlet, Google docs, Sci2U) has made good sense in relation to what we were to learn.	4.22
There was some follow-up on the mid-term evaluation.	4.21
The course lived up to the course description and the academic objectives.	4.10
The teaching was successful.	3.67
The course was structured appropriately.	3.53
During the course, I had the chance to get feedback from fellow students and/or the teacher(s).	4.14
The course activities (reading, exercises, excursions etc.) were relevant.	4.37
In my opinion, the exam form is consistent with the form of teaching and other course activities.	4.70

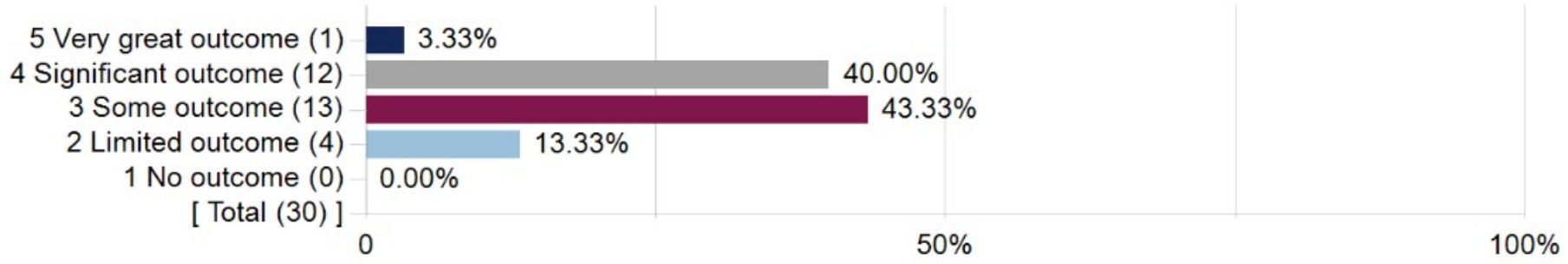
Student engagement





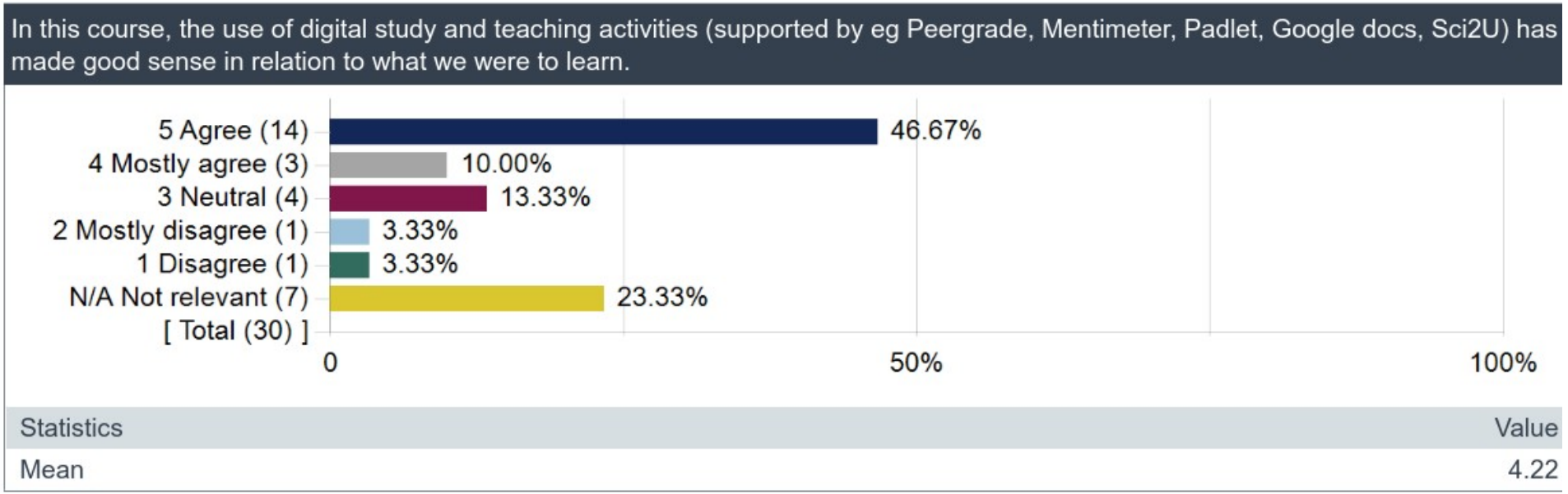
Overall outcome

I rate the overall outcome of the course as:



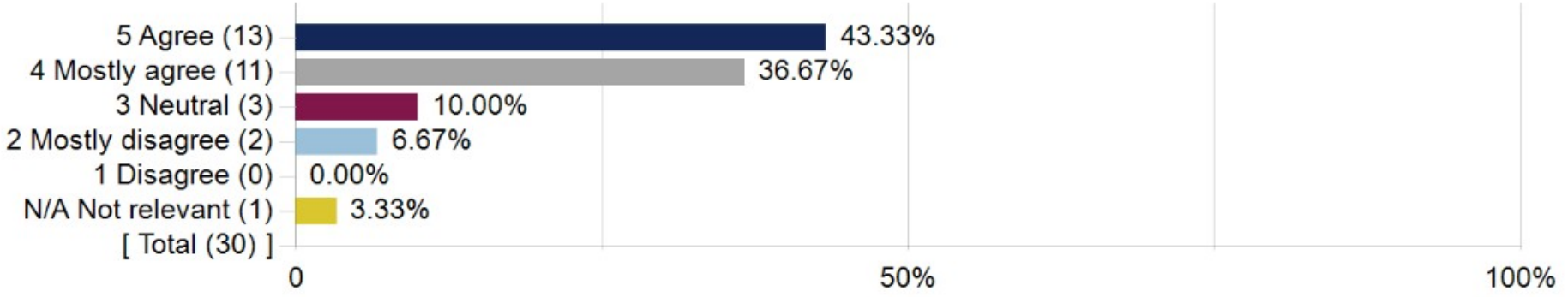
Statistics	Value
Mean	3.33

Digital teaching activities



Mid-term evaluation

There was some follow-up on the mid-term evaluation.



Statistics	Value
Mean	4.21

Mid-term evaluation – part 1

Comment
- Honestly, I felt like the midterm evaluation was presented to us quite late, so it might have been a little hard to make some radical changes throughout the remaining weeks.
- We talked about it, but changes were never made
- Lau har været virkelig god til at modtage feedback og har søgt at ændre på det, der har været muligt.
- there was follow up on some areas, but not in the areas that would've helped me more eg more in depth walk throughs of the coding aspects, more clarity in what is expected from us etc
- Der har været mere struktur og længere frist på assignments. Der er dog stadig en del lange og kringlede opgaver, der kunne være blevet forbedret efter evalueringen.
- Som jeg husker gik kritikken på for meget matematik uden mere konceptuelt grundlag og for lange opgaver. Det mener jeg ikke har ændret sig synderligt.
- Questions for the portfolio were phrased in a slightly more clear way, but they still left something to be desired.

Clear improvement	n = 1
Some improvement	n = 3
No improvement	n = 2

Mid-term evaluation – part 2

- We followed up on the feedback very well and Lau showed great attention to the issues we had. However, there were still some of the main issues that have not really been resolved fully but that is to be expected.
- There were some follow–up, but some of the major problems we talked about in the mid–term evaluation was not completely fixed (but I know it's a new course and it's difficult to make everything perfect!! i'm just happy we could have a dialogue about what did not work out). I still think there is a huge discrepancy with how much time was expected of us to use for the portfolios (i head 10 hours – i have spent minimum 25 hours for each of them...).
- Der blev fulgt op, men kan ikke mærke forskellene i praksis efter opfølgning
- Lau is good at asking for and receiving feedback
- The teacher made a sincere attempt to make changes in accordance with problems brought up in the mid–term evaluation. However, there was still a feeling that the teacher did not fully appreciate just how much many students were struggling with their learning. This may be because of the group work, which is good otherwise, but has the downside of hiding those that have almost no understanding, because it only requires one or two students per group who understand the material to complete the exercises.
- Lau is incredibly good at listening to feedback – not only at the mid–term evaluation but throughout the entire course.

Clear improvement n = 1+3 = 4
Some improvement n = 3+2 = 5
No improvement n = 2+1 = 3

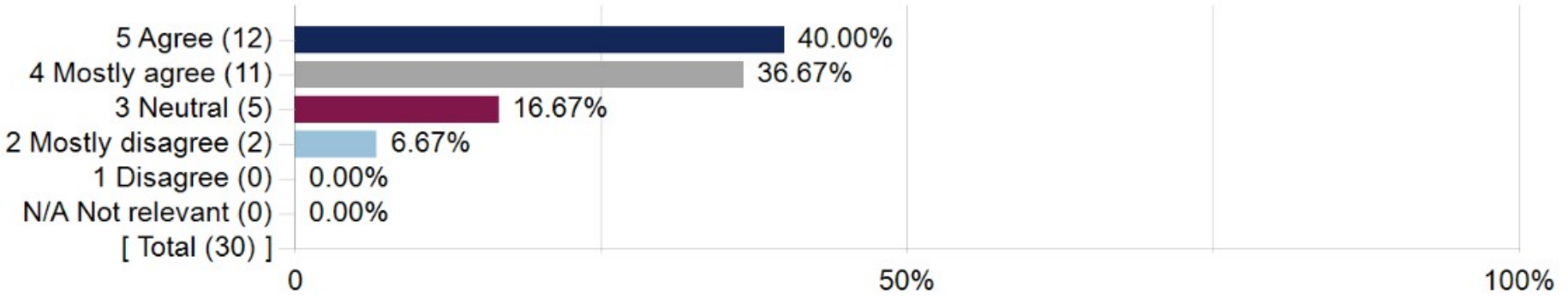
Mid-term evaluation – number of exercises

- Before mid-way evaluation
- 5 practical exercises over 6 weeks with 15 sub-exercises total
- 15/6 = 2½ exercises per week

- After mid-way evaluation
- 3 practical exercises over 5 weeks with 6 sub-exercises total
- 6/5 = 1.2 exercises per week

Description and objectives

The course lived up to the course description and the academic objectives.



Statistics	Value
Mean	4.10

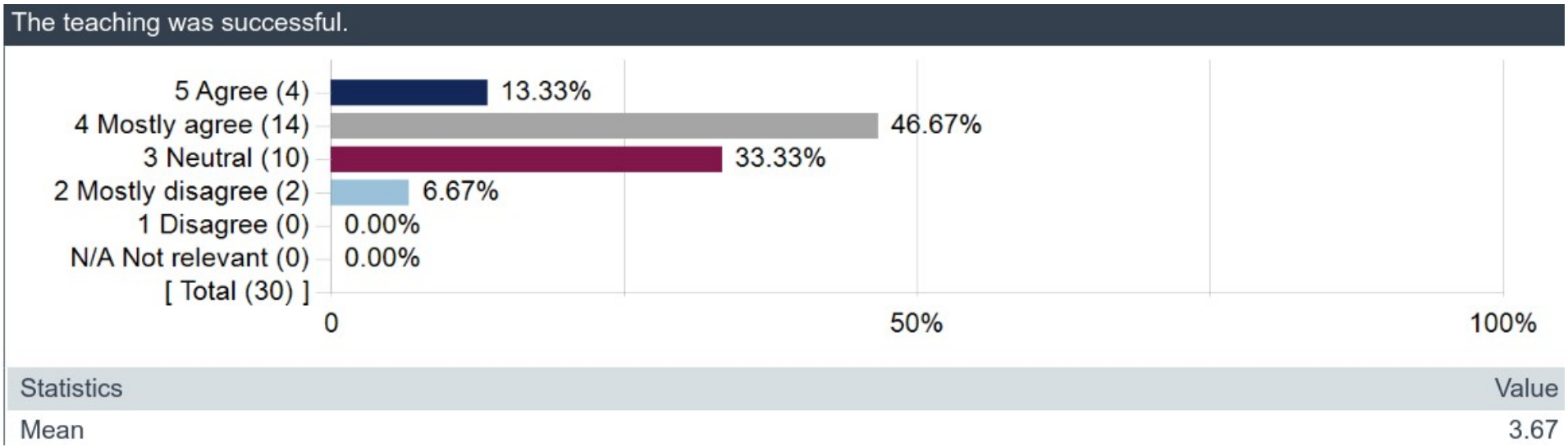
Description and objectives

Comment
- We've reached the academic goals but also went far beyond that.. I feel more confused than educated.
- Vi har fået en ret grundig forståelse af Multilevel Modeling, mens Machine Learning har været forholdsvis overfladisk (men det har været, hvad man kunne forvente af et så godt kursus i noget så komplekst).
- i think other students would find that it did but just for me i didn't get much out of it
- Kurset har ramt lidt over niveau, så det er lidt svært at evaluere hvor meget jeg egentlig har fået med. Det har sikkert levet op til de faglige mål, men fordi det har været på så højt et niveau er meget af det gået tabt for mig.
- I'm unsure how much I would be able to apply the concepts learned on my own. We had so many assignments and weren't allowed to get feedback due to them being part of the final exam portfolio. It's hard to know what I/we did correctly.
- I think the academic objectives have been too ambitious so maybe Lau have thought us everything, but I think it has been so difficult so I haven't really gotten a lot from the courses
- Har svært ved at vurdere om jeg har nået de faglige mål
- YES, but I think the course does not build well enough on top of methods 2 and I do not think this is our teacher's fault but rather a structural issue. We do not know as much math as expected from methods 2. Atleast not well enough to feel confident in applying it. It felt more like a "taste" of math than anything. We have also had a break from working with messy data which made the pre–processing part in R and Python difficult.

STRUCTURE / MISMATCH OF EXPECTATIONS

IT'S HARD (IT IS, I AGREE)

Teaching quality



Teaching quality – part 1

Comment
- Både Lau og Emil er rigtig søde!! Der er bare så meget, der går så hurtigt...
- The lectures has been loaded with a lot of new theory and meanwhile we've been learning Python, so personally I found it really hard to find a standpoint where I felt like getting the gist of any of them.
- To some extent I guess, but I couldn't follow along very well
- It has worked OK but it's sometimes hard to follow all the information in this tempo.
- Nogle gange har forelæsningerne blevet stoppet for meget af nogle meget specifikke begreber, som folk har haft svært ved at forstå. Mange af disse har ikke nødvendigvis været så relevante for den overordnede forståelse. Dog er det dejligt, at Lau har været klar til at besvare spørgsmål grundigt.
- as previously mentioned, i think the teaching worked well for others, and I think Lau is a great teacher, i think it was just a combination of the confusion from the previous course meant that knowledge he expected us to already have and therefore build on that in order to succeed, i feel many of us didn't have that knowledge and there wasn't much done to help fill that gap
- Lau er god til at clarify når det bliver spurgt.
- Sometimes a lack of concrete feedback meant that I wasn't 100% clear as to how well I am doing.
- It has been too difficult, but Lau is a very nice man and nice teacher who really tries to explain, when one asks a question
- Lau is great, but I often feel like there is too much information in each lecture; you kind of end up feeling you understand most things superficially, but don't really understand anything completely. That is a bit frustrating sometimes – I would have a hard time using many of the methods we've learn independently (in an exam project, for example), because I don't feel secure enough in using them, since it haven't understood the different methods to the core.
- Der er mange af emnerne jeg ikke forstår, og som er blevet forklaret på en måde hvor jeg ikke kan relatere dem til helheden af kurset

IT'S HARD (IT IS, I AGREE)

Teaching quality – part 2

- It is difficult to say precisely, because it also depends on the previous Methods–courses and the assumed knowledge of the students. While the teaching probably worked for those students who do not struggle with Methods, the main issue is that the teacher did not seem to appreciate just how much other students were struggling despite efforts.

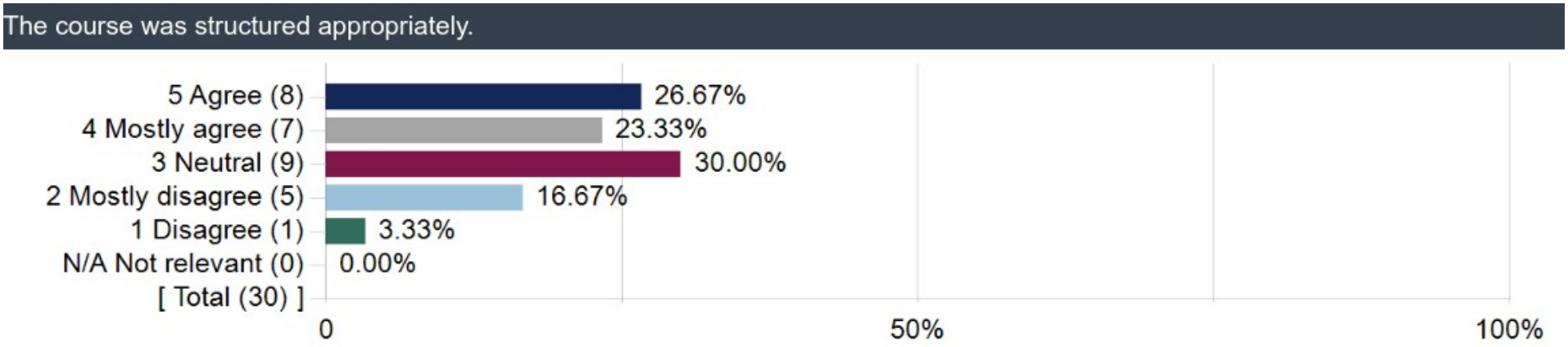
- Lau is good at explaining things

- I would have preferred more explanations of the material. In my experience, in depth explanations only occured when students asked eleborative questions – which fostered more questions and caused us to run out of time before everything was covered. This caused the confusion of having to spend next week on something that should have been covered this week and so on. When I say 'in depth' I do not mean more complex or nerdy – I would have liked some eleborate connections to what I already know. I feel like we were overestimated in what we had already learned.

- I'm not really sure how to answer this question. The teaching style itself has been good – however, there has been a 'Missing Link' between this course and the earlier courses (our actual level and the expected level) which has made it very hard to really grasp the concepts and get as much out of the lecture as I could have hoped (just small things as implicitly using new terms to already known concepts made it more confusing and harder to grasp). Also, there was also a missing link between the lecture and the exercises – many things were new in the exercises and having to do a lot of statistics without no prior knowledge was almost impossible, and it took a lot of hours just to understand many of the exercises.

STRUCTURE / MISMATCH OF EXPECTATIONS

Course structure



Statistics	Value
Mean	3.53

Course structure – part 1

Comment
- I just really miss the main thread between Methods 1–3. This semester would have been way better and more doable if I actually had a little more prior knowledge to lean on. I've kinda felt "thrown in" to this semester, not really knowing what was going on. When you start teaching students new content (ML), you've gotta start at a baby level before rushing into hard mathematical equations and assume they understand it perfectly just because they had a more mathematical second semester on zoom..
- A bit chaotic, but otherwise yes
- We could have used some introductory course to Python as this is a new coding language for us. A seminar after the autumn break would have been good. It has been decently hard to jump right into statistical and AI coding in Python.
- En rigtig fin udvikling mellem afleveringer og undervisning.
- Der var for meget content i de første uger, hvilket gjorde at læsning og forelæsninger slet ikke passede sammen. Dette var ikke hensigtsmæssigt.
- Det har været svært at vide hvad præcist det er vi skulle lære på kurset, der har manglet nogle overordnede linjer og generelt tror jeg der har manglet noget kontekst for den mængde matematik der har været. Generelt vil jeg mene der har været alt for meget fokus på formler og matematik. Jeg ved ikke om det måske kunne skyldes at det har været forventet at vi har kunne mere matematik er vi I realiteten kan. Desuden var overgangen fra R til Python fuldstændig uden introduktion og jeg vil kraftigt anbefale et eller flere python workshops/kurser til næste semester.
- I think it was really nice that we started with R, which was familiar, and then dived into Python at last.
- There are main concepts that we maybe should have known well based on our methods 2 syllabus but that wasn't the case. Huge knowledge gaps made it hard to keep up this semester and there wasn't enough review built in when concepts tied into methods 1, like Python coding. The assumption that we could use Python for one portfolio in methods 1 and then start with machine learning halfway through this semester did not benefit us or our learning.
- It would be amazing if all of the Methods teachers could sit down and tell each other what has been taught to the students so the following methods class can be structured based on this

STRUCTURE / MISMATCH OF EXPECTATIONS

PYTHON NOT INTRODUCED PROPERLY

Course structure – part 2

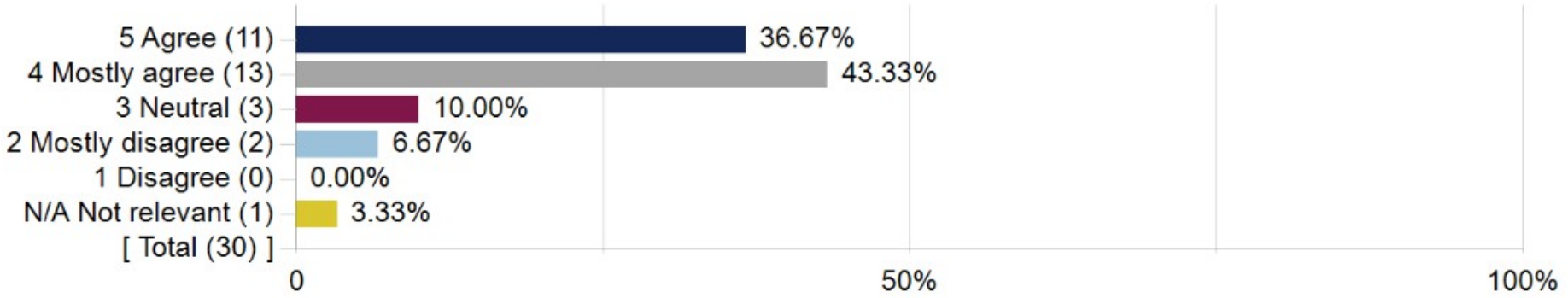
- I think there needed to be a bit better communication between the other methods courses, as to know what we have learned so far, so it is a better transition into the new methods course (this is more for the future structure of the course – e.g. maybe we could get introduced to python a bit more a bit earlier in the other methods courses, so it is not so foreign to us, when we have to start right away with doing machine learning)
- We used 2,5 semesters to learn R, and in two classes we were supposed to go from no experience with python to coding machine learning. We need a better introduction to python. The portfolios were too long and too complicated, with too little time to hand them in.
- I think the course is too cramped. Too many topics so it once again feels somewhat superficial – at least in topics which we have not previously worked with. It was much nicer to dive deep into mixed effects modelling and logistic regression which we had been exposed to in methods 1 and 2. I also lack coherence from methods 1 and 2. I am grateful for the exposure to python – this is super good!! However, I think we would benefit from a "python basics" class or two where we learn to work with 3D data in numpy and some more object–oriented programming, list comprehensions etc.
- I would have agreed to this, if we hadn't been behind schedule for some weeks. I think the syllabus made sense.
- The structure of the course material has been optimal. However, the sudden transition from Rstudio to Python was very "unnatural" as this was a new language to do statistics in.

STRUCTURE / MISMATCH OF EXPECTATIONS

PYTHON NOT INTRODUCED PROPERLY

Feedback

During the course, I had the chance to get feedback from fellow students and/or the teacher(s).



Statistics	Value
Mean	4.14

Feedback – part 1

- Lau has been so good at receiving feedback, it has been a pleasure.
- Mostly from fellow students through peer review/helping each other rout with portfolios. Not so much from the teacher
- Vi har kun modtaget meget lidt feedback angående afleveringer.
- we had code review with other students which was very helpful, but the little feedback from the teacher was always generalized to the whole class so i never really knew how to improve on the issues i was having trouble with
- Individuel respons på portfolios ville forbedre the outcome en del.
- The feedback from the professor could have sometimes been a bit more specific so that I have a chance to actually improve.
- There could have been more feedback on the exercises/portfolios – as someone mentioned in class, we do our best when working on portfolios and such and therefor does not always know whether we are wrong or right.
- I think we lacked more specific feedback in relation to assignments that are also on the final portfolio. But we had a long discussion about this and I do not feel that I am completely clueless as to how we have done on those assignments. So there has been some feedback.
- The peer review/code review is amazing!! One of my favorite things we've done. I wish we did that in all of our methods classes. Wish we'd gotten more feedback from Lau though, with students it can feel a bit like the blind leading the blind.
- Like the general feedback from the portfolios – but we are used to getting more individual feedback for the different portfolios, which we haven't really. This makes it very difficult to know if you did something wrong in some of the assignment – which is not a very nice feeling to have when you have to hand them in for an exam. So I could definetely use some more concrete feedback on the portfolios, since they were very difficult, and some of the answers to the exercises felt like "a shot in the dark".

LACK OF SPECIFIC FEEDBACK

PEER-REVIEW HELPFUL

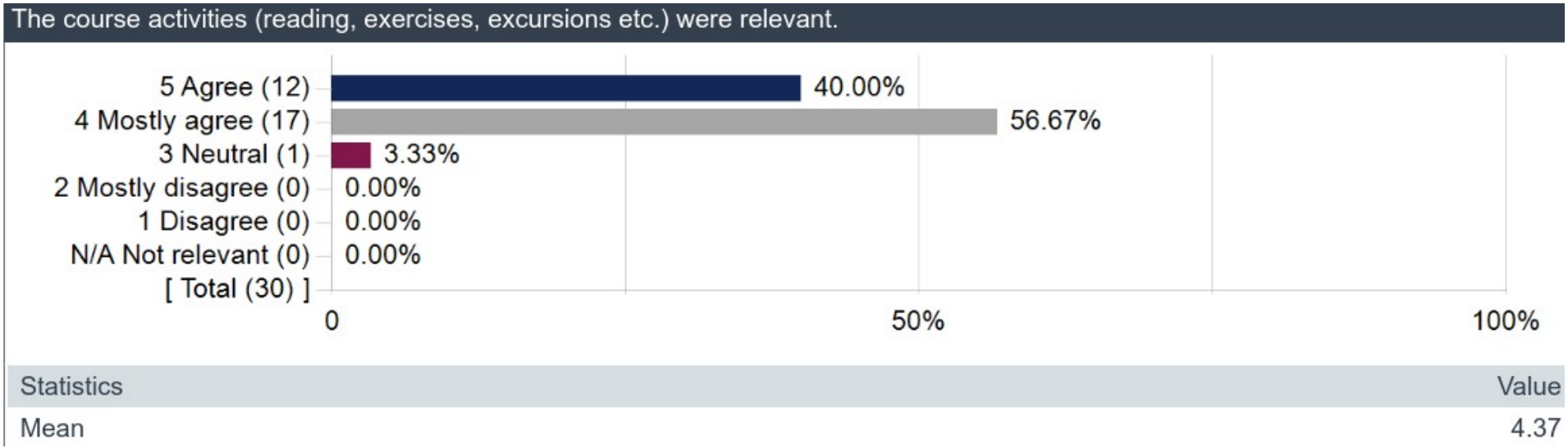
Feedback – part 1

- Det har været god dialog mellem underviser / elever, men ikke nok feedback på skriftligt arbejde
- I LOVE coding reviews.
- It feels silly, even absurd, that we are doing the portfolio exercises during the course, but cannot receive proper feedback on the exercises because "it is giving away the answers to the exam". The code review, over the shoulder –stuff was good.
- I do not think the feedback has been enough :((Especially on the assignments themselves. I know that our teacher cannot help us for the exam, but I think teachers should be able to answer questions directly related to the exam.
- We have talked about feedback in relation to the midterm evaluation. I would have liked feedback on portfolios, and I don't quite understand why we couldn't have that. In my opinion the portfolio exam form is to help you improve and learn during the semester. If we do not get feedback on our work, we might as well do and end term written exam instead. I did very much enjoy over the shoulder coding with class–mates.
- Lau has been good at answering questions.

LACK OF SPECIFIC FEEDBACK

PEER-REVIEW HELPFUL

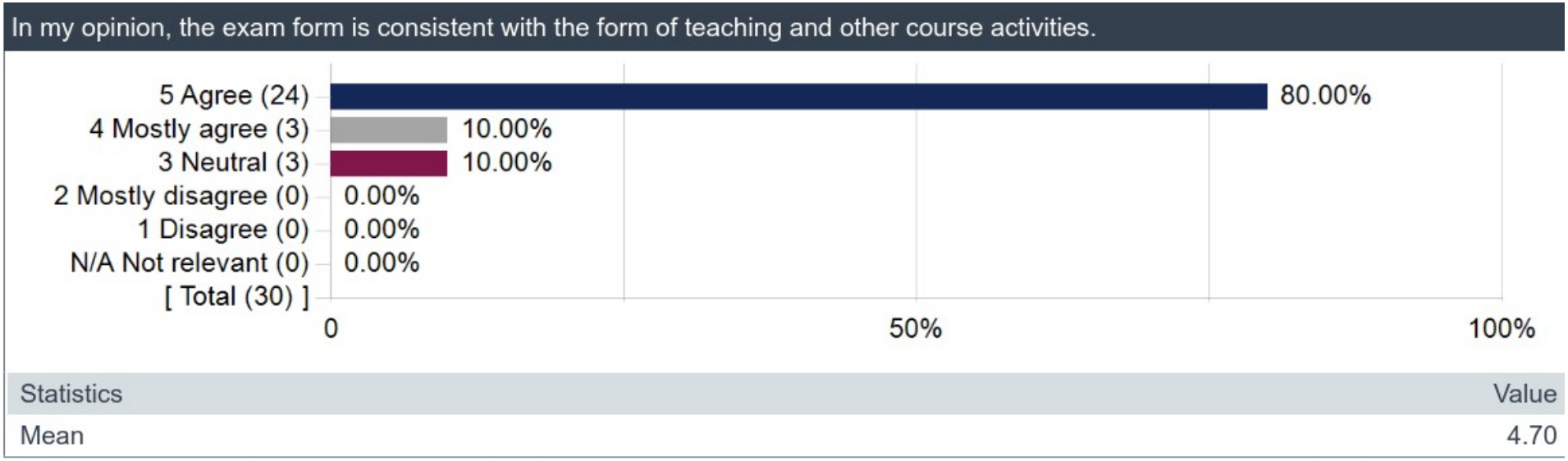
Relevance of course activities



Relevance of course activities

Comment
- They have been relevant but often, hard to finish in the time frame taken the difficulty into account.
- Til tider har læsningen været lettere irrelevant eller gentaget tidligere gennemgået materiale. Men generelt godt.
- Sometimes in our assignments we were asked to do things we hadn't really learned or had been taught how to do. We spent sooo much time reading up on these things and it got a bit frustrating.
- The readings have been good but the practical exercises have been very challenging
- I generally have found the readings to be helpful (except for the chapter on the general linear model – I Think it was by Gellman or something. Way above our level). Especially the Python Machine Learning book has been nice.
- Nogle gode tekster her og der, som formår at forklare stoffet enkelt og præcist, ex. Winter (2013).
- Would prefer to have more live coding to connect lectures and classes. The math and statistics seem to be in another sphere than the practical exercises.
- Very good book on python!!! Generally great readings and not too much but not too little. The syllabus could however benefit from a few math–related videos e.g. "today we are going to talk about PCA, so you can brush up on your eigenvectors by watching this video" so that we have a chance to catch up on the math that we may not be able to remember.
- I am a fan of the Python Machine Learning book.

Exam



Exam

Comment
- More practical teaching on coding would have been useful .. e.g. going thru code during classes and giving R–markdowns with code examples and/or assignment solutions.
- I certainly thinks this exam form makes a lot of sense. It makes it feel like an exam that is more spread out throughout the entire year and slowly accumulates. Also based on the previous methods courses, this way of learning has been extremely prosperous.
- Portfolio er rigtig godt, men mere feedback i løbet af semesteret
- I think a portfolio exam requires more feedback directly tied to the assignments. Also, would be nice to be done with our portfolio assignments at least 1 or 2 weeks prior to handing it in so that we have time for revisions...
- See other comment

Greatest contribution to learning – part 1

Comments
Holdundervisning og når min studiegruppe og andre på holdet har forsøgt at forklare mig ting. Emil er rigtig god til at forklare nogle lidt mere basic ting.
opgaver
Coding. I've really enjoyed getting hands on with coding again.
getting feedback
At lave portfolios og exercises, sådan man selv får hænderne i kodning.
Peer review, coding together with others
Portfolio opgaverne har været udfordrende, hvilket har gjort at det har været nødvendigt at tale meget om dem (f.eks. i studiegruppen) – og det synes jeg har givet et godt udbytte. Derudover synes jeg, at når vi har haft class, så har man kunne bruge nogle af de lidt abstrakte ting fra lecture, også gav det mere mening.
It has been nice to learn more about Github and other software.
The texts we had to read were really good, as they explained the material in a clear and concise way. It really helped that the texts focused on the conceptual ideas of machine learning/statistics instead of getting bogged down in heavy math. It is necessary to get the birds–eye view that a (mostly) purely conceptual description offers, which then allows the student to dig down into the mathematics afterwards.
Helt klart afleveringerne. Jeg har personligt haft det fint med kodningsdelen, så det har været rart at få en dybere forståelse af det rent metodisk teoretiske.
the first peer code review
I personally learn most when I get my hands dirty, so coding exercises have contributed a lot to my learning.
Python machine learning began

CODING AND EXERCISES

PEER-REVIEW HELPFUL

Greatest contribution to learning – part 2

It was very helpful that we were able to ask for assistance from both the professor and the instructor, both in–person and online (using Cryptpad).
Portfolios and class exercises. Lectures were really good as well but a lot the level was increased significantly this semester so it's been a bit hard to follow everything.
Definitely the exercises. They were very difficult (at some times too difficult) and time consuming. However, I think it's nice that they were ambitious and really challenged us. Also, working with e.g. brain data was something that was a great learning experience.
Code review with other students and doing the portfolios in groups. I also enjoyed the cryptpad a lot because we can all learn from each others questions.
Classes most definitely
Even though they have been very difficult, I think the portfolios and exercises have given me a lot of experience with both R and Python, which is really nice – I feel like a have gotten a lot better at coding!!
Selvstudie (litteraturen var god) og afleveringer.
Exercises, dialog med medstuderende og Emil's simple forklaringer
Sitting in classes and doing code with either Lau or Emil being there to help out/explain something, if I would get stuck. That's something that I really appreciate, having someone readily available to come and help me out and explain it conceptually to me, so I don't have to spend a lot of time on the internet/youtube, trying to find the right information.
Practical exercises and portfolios. Struggling so much is annoying and unnecessary, but practical tasks sticks better than listening to lectures and reading.
Working on the portfolio with my study group and trying to explain the code to each other.
Python, practical exercises, introduction to Github (thank you!!). Great lectures. Nice that we can ask Lau on mail and crypt pad for questions.
Doing the portfolios with my study group.
Coding exercises and examples from datasets with "non–abstract" so we were able to follow the numbers and code.

CODING AND EXERCISES

PEER-REVIEW HELPFUL

Course improvement – part 1

Der er sindssyg dårlig sammenhæng mellem methods–kurserne, hvis I spørger mig, og det er tydligt at mærke, at underviserne ikke rigtig ved, hvad vi kan fra tidligere kurser... Det gør, at der nogle gange bliver ramt meget ved siden af, hvad angår sværhedsgrad (og primært til den for svære side). Så hvis man har svært ved methods (skyldig), så er det meget nemt, at blive tabt. I methods 3 har vi brugt RStudio i første halvdel og Python i anden halvdel. Problemet er bare, at vi ikke har noget træning i Python (ud over én uge på første semester i PsychoPy), men det forventes åbenbart, at vi bare kan det. Det forstår jeg ikke, hvordan kan være en selvfølge? Derfor er et konkret forslag til at gøre kurset bedre at lave en Python–workshop!!! Jeg deltager meget gerne, hvis der bliver lavet en for dem næste år.
Ditch the repetition of R in the begning and start with ML straight away. Introduction to multi–layer NN, we could move a little faster and by starting with ML straight away we could reach some ML algorithmns that are actually useful and not just good for the conceptual understanding.
I don't even know where to start.
shorter exercises
At vi havde et grundlæggende python kursus, som lærer os at bruge python sproget i forhold til statistik.
Less maths/make it clear that the maths is optional for those who want to know more about it. Portfolios should be shorter, clearer and more concise + with longer deadlines. Give introduction to Python instead of just switching to it without telling students the basics of the language (I am aware that we were given a document with basics, but it's not very intuitive to just begin programming in another language by yourself) Better links with previous methods courses Better links with other courses in this semester. More group work/class work on portfolios and assignments (often the exercises were not covered in lectures/classes so we had to google everything ourselves to figure out how to do things) More feedback on portfolios/assignments. Teacher should be present in all practical classes
Måske indskrænke antallet af emner, så der er mere tid til at få dybere forståelse af færre emner

STRUCTURE / MISMATCH OF EXPECTATIONS

PYTHON NOT INTRODUCED PROPERLY

LACK OF SPECIFIC FEEDBACK

Course improvement – part 2

I put some already, but:
– Python seminar / workshop or just more introductory teaching for Python coding.
– More hands–on coding in classes like in first semester.
– Coding examples, going thru assignments?

The portfolio assignments need to be more interpretable in terms of what the student needs to do. More specifically, we were often handed portfolio assignments with exercises that required knowledge of a neuroscience experiment (or something similar) of which most of the students were ignorant. Furthermore, we were asked to do Machine Learning in Python, which was a bit of a challenge given that most students have not had a thorough introduction to Python – here I'm thinking of the fundamental ideas such as arrays, dataframes, syntax of the language etc.

Hold fast i afleveringernes sværhedsgrad. Det får folk til at bruge tid og kræfter på materialet.
Desuden ville et Python kursus nok gøre godt for mange, samt et kort brush–up af relevant matematik (nok primært lineær algebra).

do a lot more of a walkthrough with python, and in general have exercises that guided us on what to do, as every class it felt that every assignment was for the exam in that we had no guidance unless we kept on asking them questions every couple minutes, some text in the script explaining to use a particular kind of code and what for would be very helpful in helping me learn, or at least a solution markdown to be able to refer to as we have had in the previous courses, as I felt there was no learning curve, you either knew what you were doing or you didn't and if you didn't it was very hard to actually do anything on your own

The teacher repeatedly overestimated our skills in statistics/programming. Generally, the skill level required to do the portfolios/exercises was too high (at least for our study group). We didn't do much coding in our 2nd semester which probably meant that Lau expected us to be somewhere where we weren't. So perhaps for the future, the teachers of 2nd and 3rd semesters should have a more thorough talk about what the students have learned and have not.

This was mostly an issue in the first half of the course, but the readings could have been slightly more relevant to what was being discussed in class. The feedback from the professor could have sometimes been a bit more specific so that I have a chance to actually improve. Sometimes a lack of concrete feedback meant that I wasn't 100% clear as to how well I am doing.

Double the amount of lectures and take us more gently through the tough maths and concepts. I believe that many of us did not pick CogSci for the coding part and are therefore not as inclined to pick up on everything with the current pace

STRUCTURE / MISMATCH OF EXPECTATIONS

PYTHON NOT INTRODUCED PROPERLY

LACK OF SPECIFIC FEEDBACK

Course improvement – part 3

I think an introduction to more general Python would be advantageous. I personally have some experience with the language, but for those who haven't it seemed to be quite the challenge to just start out with difficult issues that would have been hard to work with even in a familiar language such as R. However, that is not to say that the addition of using Python in the course was a bad thing whatsoever. There should just be an understanding for the fact that the languages might not be so similar to people, who have only learned to use one programming language like R. I think this realisation of the similarities only comes with the experience of more languages.

More python or a python workshop so we don't spend so much time getting a command of the basics which I still don't feel I have.

More review from previous semesters. Possibly some sort of optional math help outside of class. This is super complex stuff when we were only required to have Math B to get into the program.

Correct the syllabus such that it fits better with what we've already been taught. Make it less ambitious –you've lost at least half of the class during this semester and during each lecture. I don't even know which questions to ask.

Less information, more feedback on the portfolios! And it would be nice with a brush–up python course – we haven't used it since the first semester (which was also only for psychopy) and it is difficult getting into using NumPy etc (just understanding the data and 3D arrays for example).

But mostly, I think there needs to be a more common understand of how long the portfolios should take – because right now, it feels very misjudged, being expected to use 10 hours, when actually spending 25+.

More focus on application and teaching by examples. Also, relate topics more to each other to get a feeling of how they depend on each other!

Sometimes in the lectures, there is a lot of formulas and math, than can be a bit hard to get your head around/grasp conceptually – I would be nice, that when being introduced to a new concept, then trying to explain it with a metaphor or something, so everyone gets a basic understanding of the concept, and from there, you can move on to the more theoretically heavy parts. Just to get a good foundation to build upon, instead of the formulas being very abstract and therefore hard to understand conceptually.
Also, I personally love StatQuest, which have helped me understand the topics of this course better – maybe you could suggest people watching some of their relevant videos, if they're having a hard time to understand what's happening.

Write a methods book – all of the methods teachers together. The disconnect between methods 1, 2 and 3 is extremely problematic, and has at one point or another had every person in my study group consider dropping out of cogsci.
The math is too complicated – we could easily spend a semester on just a few of the concepts, either cut down on the math or make it clear when it is optional.

Less cramped. More coherence with methods 1 and 2. MUCh shorter portfolios (we have spent like +20 hours as a study group on most of them despite feeling quite confident in coding). The length of the portfolios make it stressful and I feel that we learn less since we have to rush through certain parts.

Methods 2 simply does not make us confident enough in math to be able to follow along to everything happening in methods 3. Giving us videos or reminding us that we are going to talk about design matrices, eigenvectors etc. would help.

STRUCTURE / MISMATCH OF EXPECTATIONS

PYTHON NOT INTRODUCED PROPERLY

LACK OF SPECIFIC FEEDBACK

PYTHON NOT INTRODUCED PROPERLY

A six-week Python workshop will be part of 2nd semester in the future

LACK OF SPECIFIC FEEDBACK

In the regulations, there are only 30 minutes set aside for feedback per student. May be something to discuss in *Uddannelsesnævnet*

STRUCTURE / MISMATCH OF EXPECTATIONS

We need a good discussion of this in *Uddannelsesnævnet*

IT’S HARD (IT IS, I AGREE)

I think this correlates with the structure and mismatch of expectations

CODING AND EXERCISES

Despite the challenging and long exercises, they were appreciated

PEER-REVIEW

Seems to have been a success

What were our initial goals?

Overall idea of the course

FROM FIRST LECTURE

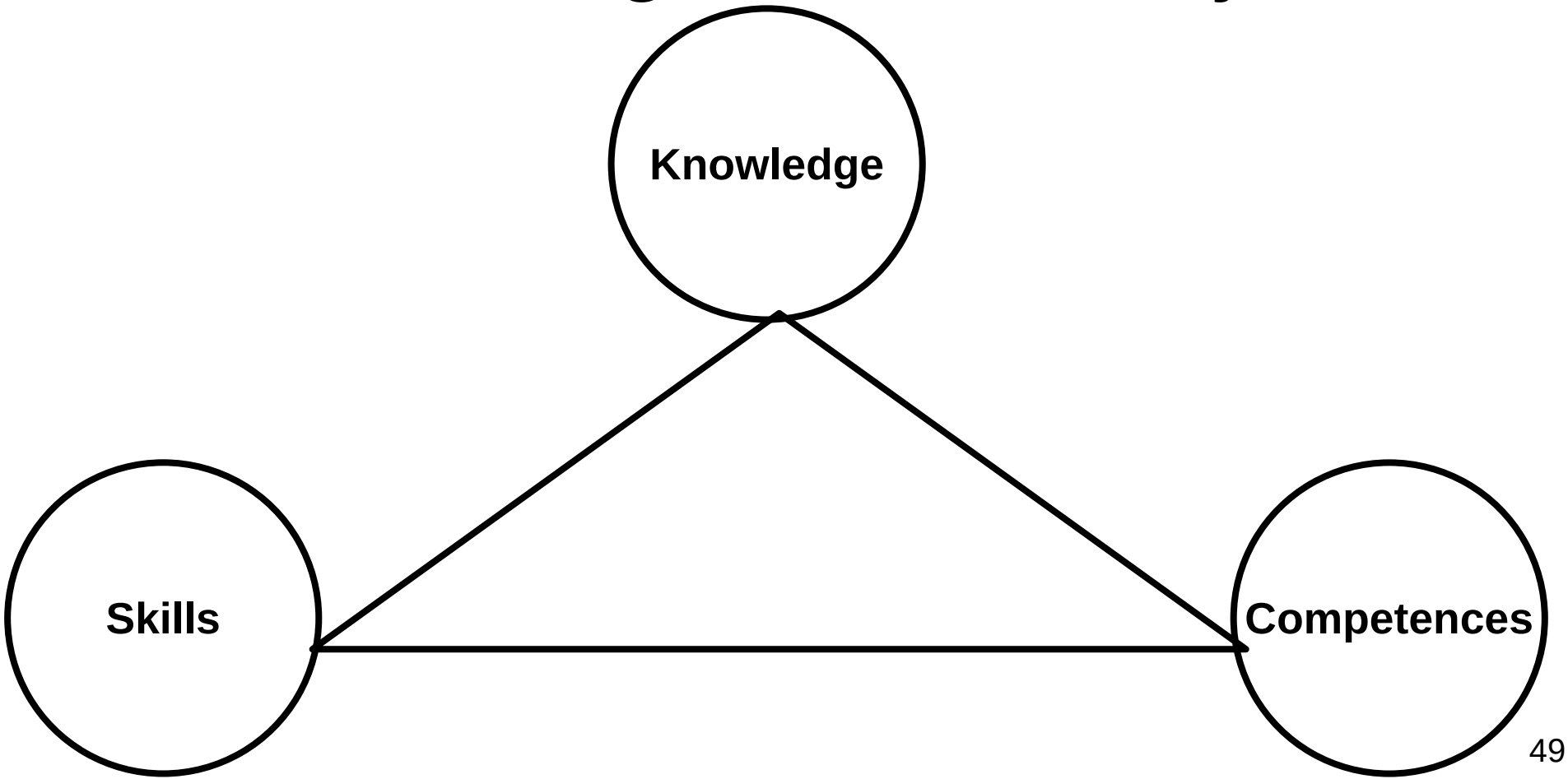
- **Master complex data:** build models that work despite missing data entries, and which model the nested and hierarchical structures of the data
- **Create generalisable models:** do model comparisons and learn the difference between *explanation* and *prediction*
- **Reduce the dimension of data:** extract the meaningful dimensions of data and separate it from the noise

Overall idea of the course

FROM FIRST LECTURE

- The aim is that you learn the practical skills (not just the theory), such that you yourself can build and evaluate models in *R* and *Python.*
- In a wider sense, the idea is that you should be able to critically assess models and understand their advantages and limitations.

Academic regulations – objectives



Academic regulations – objectives

Knowledge:

- demonstrate understanding of statistical techniques relying on the Generalised Linear Model

- demonstrate understanding of hierarchical modeling methods

- demonstrate understanding of basic machine learning concepts.

Academic regulations – KNOWLEDGE

You should thus be able to answer questions like:

What is logistic regression?

What are some naturally occurring hierarchies within research data?

What is cross-validation?

What is a mixed-effects model?

Academic regulations – objectives

Skills:

- build and evaluate models of hierarchically structured data

- integrate machine learning procedures in data analysis

- communicate analysis processes, results and interpretation.

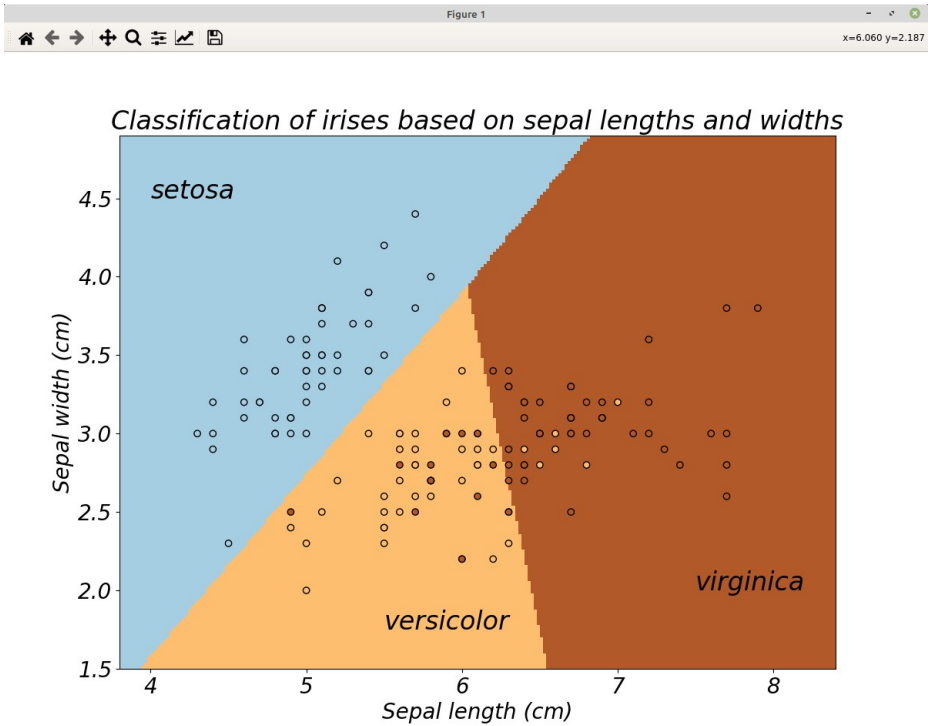
Academic regulations – SKILLS

build and evaluate models of hierarchically structured data

```
Linear mixed model fit by REML ['lmerMod']
Formula: Height ~ Gender + (1 | Family)
   Data: height
```

Academic regulations – SKILLS

integrate machine learning procedures in data analysis



Academic regulations – SKILLS

communicate analysis processes, results and interpretation

(Andersen et al., 2019)

Mixed model analyses (McCulloch & Neuhaus, 2005) were applied to investigate how top-down expectations (No. of Possible Targets) affected subjective experience and objective performance. We performed model comparisons between models that did or did not include the relevant fixed effects and interactions to find the best compromise between an explanatory and a parsimonious model. This was done using the log-likelihood ratio between two models because this ratio approximates a chi-square distribution. A chi-square test can thus be used to assess whether two models differ significantly, where the test statistic is the log-likelihood-ratio and the degrees of freedom is the difference in free parameters of the two models.

Academic regulations – objectives

Competences:

- independently decide on data analysis methods, given a data set and a research question

- justify decision making when pre-processing messy data for data analysis.

Academic regulations – COMPETENCES

independently decide on data analysis methods, given a data set and a research question

justify decision making when pre-processing messy data for data analysis.

EXAMPLE QUESTION: Can native Danish speakers tell apart soft d's (ð) and l's if they are embedded in English speech, e.g. maðfunction/malfunction

EXAMPLE DATA: Dataset with response times and discrimination responses

Exam; portfolio

- Ongoing assignments to be solved in small groups
  - Will be done in *R Markdown* (.Rmd)
- Final portfolio:
  - Revised assignments, handed in as short reports with reproducible code on ~~GitHub~~ the Digital Exam system
    - Assignment 1: Using mixed effects modelling to model hierarchical data (Winter & Grawunder, 2012)
    - Assignment 2: Mixed effects modelling of response times, response counts, and accuracy (Andersen et al., 2019)
    - Assignment 3: Using logistic regression to classify subjective experience from brain data
    - Assignment 4: Dimensionality reduction, finding the signal among the noise

The course plan

Week 1: *Introduction and Why are we here?*, September 14 & 15

Instructor sessions: *Setting up* R *and* Python *and recollection of the general linear model*

Week 2: *Linear Mixed Effects Models*, September 21 & 22

Instructor sessions: *Modelling random effects – and how do they differ from fixed effects?*

Week 3: *Generalized Linear Mixed Effects Models*, September 28 & 29

Instructor sessions: *What to do when the response variable is not continuous?*

Week 4: *Explanation and Prediction*, October 5 & 6

Instructor sessions: *Why are good explanations sometimes bad?*

Week 5: *Evaluating and comparing models,* October 12 & 13

Instructor sessions: *How do we assess how models compare to one another?*

Week 6: *Mid-way evaluation and Machine Learning Intro,* November 2 & 3

Instructor sessions: *Moving the goal away from explanation towards prediction and getting Python running*

Week 7: *Linear regression revisited (machine learning)*, November 9 & 10

Instructor sessions: *How to constrain our models to make them more predictive*

Week 8: *Logistic regression (machine learning)*, November 16 & 17

Instructor sessions: *Categorizing responses based on informed guesses*

Week 9: *Dimensionality Reduction, Principled Component Analysis (PCA)*, November 23 & 24

Instructor sessions: *What to do with very rich data?*

Week 10: *Organising and preprocessing messy data,* November 30 and December 1

Instructor sessions: *How to clean up?*

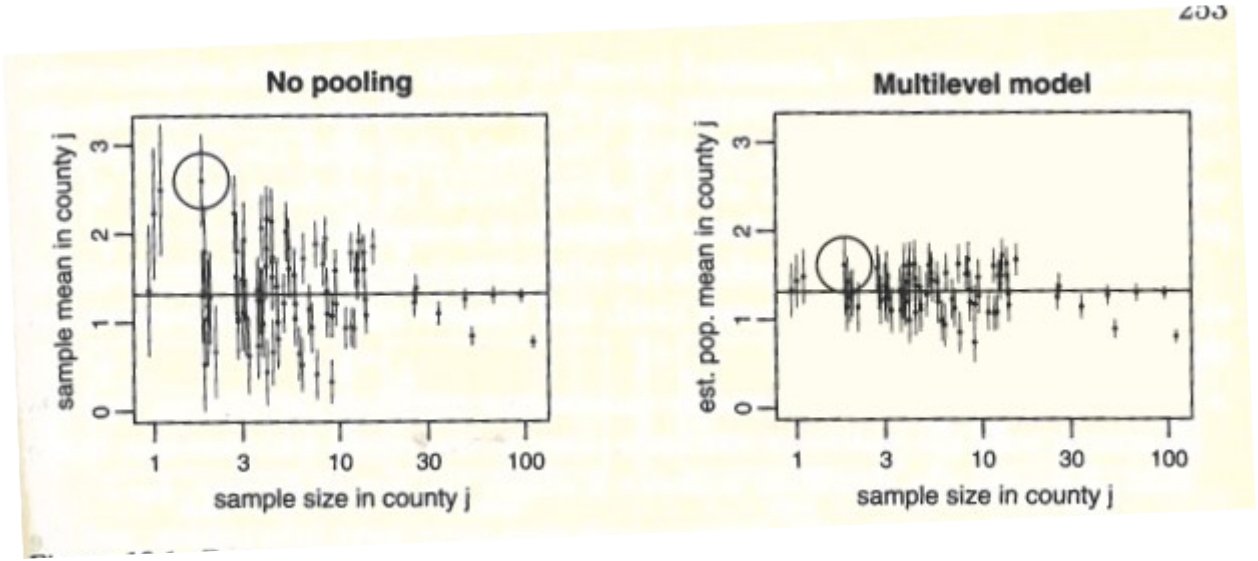
Week 11: *Final evaluation and wrap-up of course*, December 7 & 8

Instructor sessions: *Ask anything!*

What have you learnt?

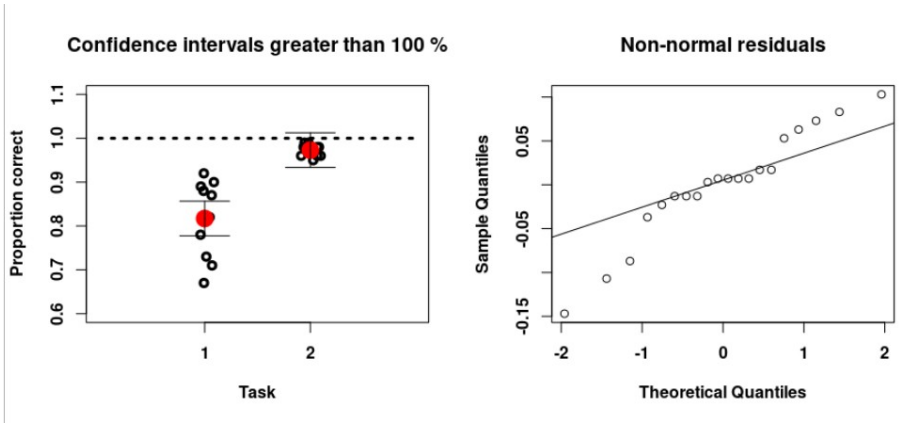
(Generalized) Linear Mixed Effects Models

Using all the data in a sensible way



(Gelman and Hill, 2006)

Using all the data in a sensible way




```r
## single-level

means.task.1 <- numeric(n.subjects)
means.task.2 <- numeric(n.subjects)
for(subject.index in 1:n.subjects)
{
    means.task.1[subject.index] <-
        mean(data$accuracy[data$subject == subject.index & data$task == 1])
    means.task.2[subject.index] <-
        mean(data$accuracy[data$subject == subject.index & data$task == 2])
}

mean.data <- data.frame('subject'=factor(rep(1:n.subjects, 2)),
                        'mean.accuracy'=c(means.task.1, means.task.2),
                        'task'=factor(rep(c(1, 2), each=n.subjects)))

model <- lm(mean.accuracy ~ task + 0, data=mean.data)
```

Introducing multilevel modelling – modelling individual slopes and intercepts

Level 1 $$y_{level} = \alpha_{level} x_{level} + \beta_{level} + \epsilon_{level}$$

Level 2 $$\alpha_{level} = \gamma_1 + S_{\alpha,level}$$
$$\beta_{level} = \gamma_2 + S_{\beta,level}$$

Variance components $$\langle S_{\alpha,cyl}, S_{\beta,cyl} \rangle \sim N(\langle 0,0 \rangle, \Sigma)$$

$$\Sigma = \begin{pmatrix} \tau_\alpha^2 & \rho\tau_\alpha\tau_\beta \\ \rho\tau_\alpha\tau_\beta & \tau_\beta^2 \end{pmatrix}$$

$$\epsilon_{level} \sim N(0, \sigma^2)$$

Examples of levels:
Subjects, Schools, Car brands
Any factorial variable, really

```
library(lme4)
multilevel.linear <- lmer(continuous ~ predictors + (predictors | subject),
                          data=data)
multilevel.logistic <- glmer(accuracy ~ predictors + (predictors | subject),
                             data=data, family='binomial')

multilevel.poisson <- glmer(counts ~ predictors + (predictors | subject),
                            data=data, family='poisson')
```

Also called
Fixed and random effects

- Fixed effects
  - exhaust the population
  - express average effects
  - can be categorical or continuous
- Random effects
  - sample the population
  - express individual effects
  - has to be categorical

Generalized linear models

1) A data vector: $y = (y_1, ..., y_n)$
2) Predictors: X and coefficients β, forming a linear predictor $X\beta$
3) A *link function* g: yielding a vector of transformed data $\hat{y} = g^{-1}(X\beta)$ that are used to model the data
4) A data distribution: $p(y|\hat{y})$

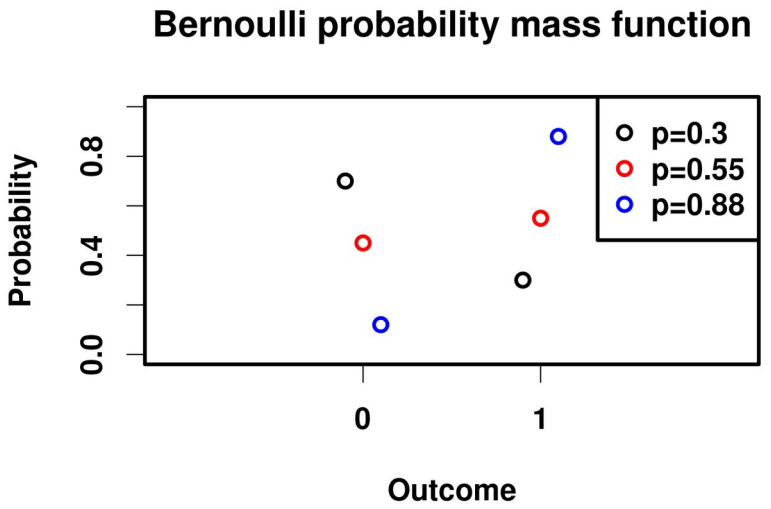
$(X\beta = \beta_0 + X_1\beta_1 + ... + X_k\beta_k)$

(Gelman and Hill, 2006, Chapter 6)

4) A data distribution: $p(y \mid \hat{y})$

$$\mathbf{PMF_{Bernouilli}} = p^k q^{1-k}$$

```
dbernoul <- function(p, k) p^k * (1 - p)^(1 - k)

x <- c(0, 1)
pmf <- dbernoul(p=0.3, x)
par(font.lab=2, font.axis=2)
plot(x -0.1, pmf, xaxt='n', xlab='Outcome', ylab='Probability', ylim=c(0,
1),
     xlim=c(-1.1, 2.1), main='Bernoulli probability mass function')
axis(side=1, at=c(0, 1), labels=c(0, 1))

pmf <- dbernoul(p=0.55, x)
points(x, pmf, col='red')

pmf <- dbernoul(p=0.88, x)
points(x + 0.1, pmf, col='blue')

legend('topright', pch=1, col=c('black', 'red', 'blue'),
       legend=c('p=0.3', 'p=0.55', 'p=0.88'), text.font=2)
```



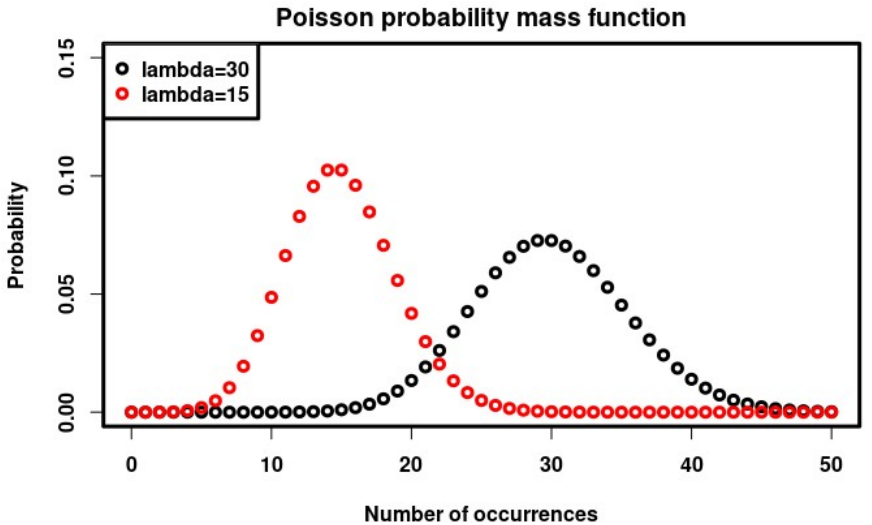
4) A data distribution: $p(y \mid \hat{y})$

Poisson(λ)

$\lambda \in (0, \infty)$: the Expected value (rate)

$k \in N_0$

$$PMF = \frac{\lambda^k e^{-\lambda}}{k!}$$



Nomenclature

- General Linear Model (single level modelling)
  - $y = X\beta + \epsilon$
- Generalized Linear Model (single level modelling)
  - $y = X\beta + \epsilon$ – but now with link functions and other data distributions besides the Gaussian one
- General Linear Mixed Model (multilevel modelling)
  - $y = X\beta + Zu + \epsilon$
- Generalized Linear Mixed Model (multilevel modelling)
  - $y = X\beta + Zu + \epsilon$ – but now with link functions and other data distributions besides the Gaussian one

```
library(lme4)
multilevel.linear <- lmer(continuous ~ predictors + (predictors | subject),
                          data=data)
multilevel.logistic <- glmer(accuracy ~ predictors + (predictors | subject),
                             data=data, family='binomial')

multilevel.poisson <- glmer(counts ~ predictors + (predictors | subject),
                            data=data, family='poisson')
```

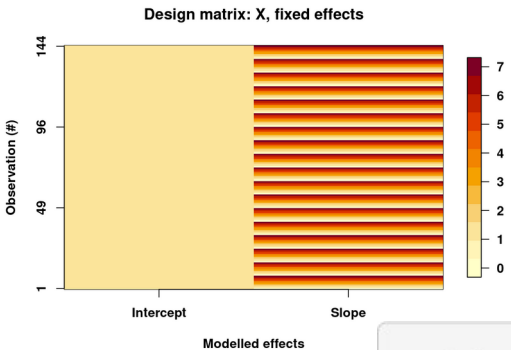
The general linear mixed model (GLMM)

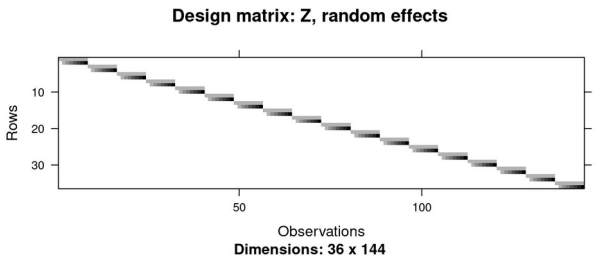
$$y = \boldsymbol{X}\beta + \boldsymbol{Z}u + \epsilon$$

$y : N\ x\ 1$ column vector
$X : N\ x\ p$ matrix of p predictor variables
β: unknown $p\ x\ 1$ column vector of the first level regression coefficients
$Z : N\ x\ q$ design matrix for the q random effects
u: unknown $q\ x\ 1$ column vector of the second-level effects
ϵ: $N\ x\ 1$ column vector of the residuals

The general linear mixed model (GLMM)

$$y = \boldsymbol{X}\beta + \boldsymbol{Z}u + \epsilon$$






```
model <- lmer(...)

Xt <- t(getME(model, 'X'))
Zt <- getME(model, 'Zt')
```

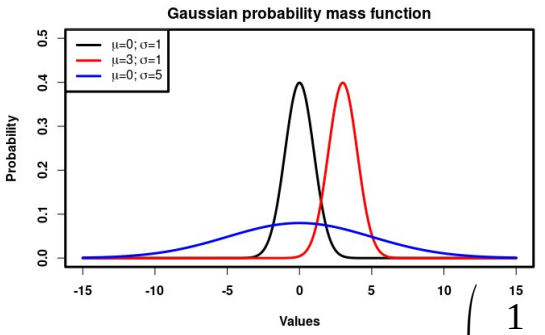
Fitting models

Likelihood function

GAUSSIAN

probability distribution function

$$\mathbf{PDF} = \left(2\pi\sigma^2\right)^{-1/2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



$$\textbf{Likelihood function}: L\left(\sigma^2, \epsilon\right) = \left(2\pi\sigma^2\right)^{-N/2} e^{-\left(\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i - x_i\beta\right)^2\right)}$$

$$\textbf{Log-Likelihood function}: l\left(\sigma^2, \epsilon\right) = \log\left(L\right)$$

y : dependent variable
$\boldsymbol{X}\beta$: linear predictor
$y - \boldsymbol{X}\beta = \epsilon$: residuals
N : number of observations

```
model <- lmer(...)

l <- logLik(model)
```

Likelihood function

BERNOUILLI (LOGISTIC)

the product of all probabilities of an observation given a probability

$$L(p) = \prod_{i=1}^{N} p^{y_i}(1-p)^{(1-y_i)}$$

$$l(p) = \sum_{i=1}^{N} y_i \ln(p) + (1-y_i)\ln(1-p)$$

```
model <- glmer(...)

l <- logLik(model)
```

Likelihood function

POISSON

$$L(\lambda) = \prod_{i=1}^{N} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$l(\lambda) = -n\lambda + \ln(\lambda) \sum_{i=1}^{N} x_j - \sum_{i=1}^{N} \ln(x_i!)$$

```
model <- glmer(...)

l <- logLik(model)
```

Assessing models

R^2; adjusted

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n} \left(y_i - \mu_y \right)^2}$$

$$\bar{R}^2 = 1 - \left(1 - R^2\right) \frac{n-1}{n-p-1}$$

n : number of observations

p : number of predictors beyond the constant term

Likelihood-ratio test

$$LR = -2\left(l(\theta_2) - l(\theta_1)\right)$$

```
anova(model.ranint, model.ranslope.and.int)
```

```
## Data: sleepstudy
## Models:
## model.ranint: Reaction ~ days_deprived + (1 | Subject)
## model.ranslope.and.int: Reaction ~ days_deprived + (days_deprived | Subject)
##                        Df    AIC    BIC  logLik deviance  Chisq Chi Df
## model.ranint            4 1446.5 1458.4 -719.25   1438.5
## model.ranslope.and.int  6 1425.2 1443.0 -706.58   1413.2 25.332      2
##                        Pr(>Chisq)
## model.ranint
## model.ranslope.and.int  3.156e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Information criteria

$$deviance = -2l(\hat{\theta})$$

$$AIC = deviance + 2k$$

k : number of predictors

```
model <- lmer(...)

information.criterion.1 <- AIC(model)
information.criterion.2 <- BIC(model)
```

When we add *k* predictors that are pure noise, deviance is reduced by an amount corresponding to a χ^2 distribution with *k* degrees of freedom.

(Gelman and Hill, 2006)

"On average, a predictor needs to reduce the deviance by 2 to in order to improve the fit to new data"

(Gelman and Hill, 2006, p. 525)

MULTILEVEL MODELS

You have learnt to use all the data in a sensible way and to assess fitted models

BOG

Python machine learning : unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics

Sebastian Raschka author; Randal S Olson author of foreword

2015; 1st edition

Tilgængelig online >

Changing focus from error minimisation to prediction accuracy maximisation (Machine Learning)

What constitutes a good model?

Remember Emil’s slides from week 03

- Accurate estimation of the underlying parameters of the population distribution

EXPLANATION

- Generalisation to new data

PREDICTION

Partitioning error into *bias* and *variance*

$$E[(y_0-\hat{f}(x_0))^2]=bias(\hat{f}(x_0))^2+var(\hat{f}(x_0))+\sigma^2$$

here you have lots of variance, so its hard to predict where the next arrow will hit

$$bias(\hat{f}(x_0))=E[\hat{f}(x_0)]-f(x_0)$$

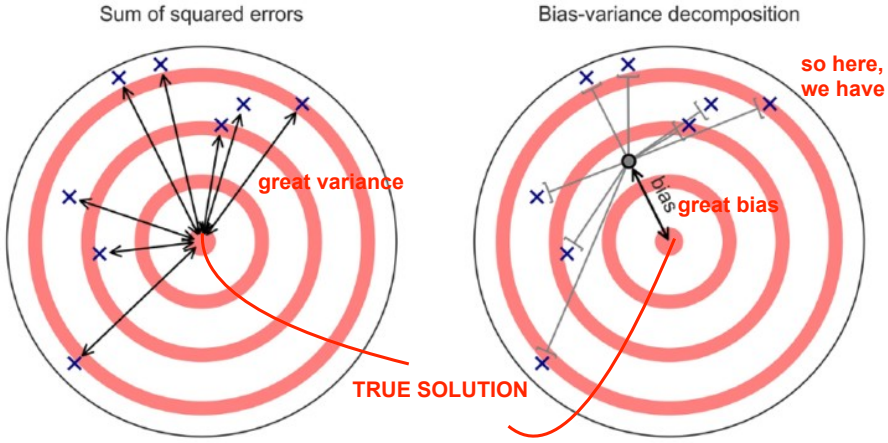


Fig. 3. Schematic illustration of the bias-variance decomposition. (Left) Under the classical error model, prediction error is defined as the sum of squared differences between true scores and observed scores (black lines). (Right) The bias-variance decomposition partitions the total sum of squared errors into two separate components: a bias term that captures a model's systematic tendency to deviate from the true scores in a predictable way (black line) and a variance term that represents the deviations of the individual observations from the model's expected prediction (gray lines).

so here, you hit it sligthly wrong (not the center), but with bias, we have a better idea of where the next arrow will hit.

we move away from the true solutoion, by introducing bias, and therefore decreaseing variance:
THIS GIVES A BETTER PREDICTION , FOR WHERE THE NEXT ARROW WILL HIT

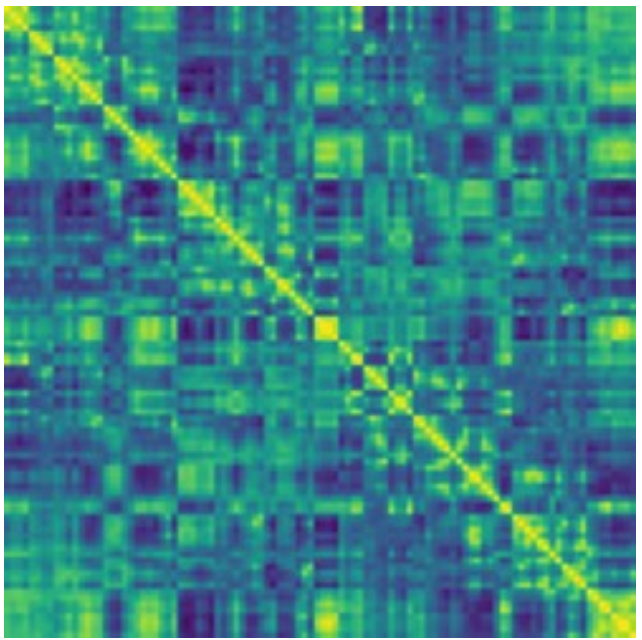
```
bias_1 = np.mean(predictions_1_and_2[0]) - y_true[x == 3]
bias_2 = np.mean(predictions_1_and_2[1]) - y_true[x == 3]
bias_5 = np.mean(predictions_2_and_5[1]) - y_true[x == 3]

# Using the formula above
variance_1 = np.var(predictions_1_and_2[0])
variance_2 = np.var(predictions_1_and_2[1])
variance_5 = np.var(predictions_2_and_5[1])

epsilon = np.random.normal(scale=5, size=100)

MSE_1 = bias_1**2 + variance_1 + np.var(epsilon)
MSE_2 = bias_2**2 + variance_2 + np.var(epsilon)
MSE_5 = bias_5**2 + variance_5 + np.var(epsilon)
```

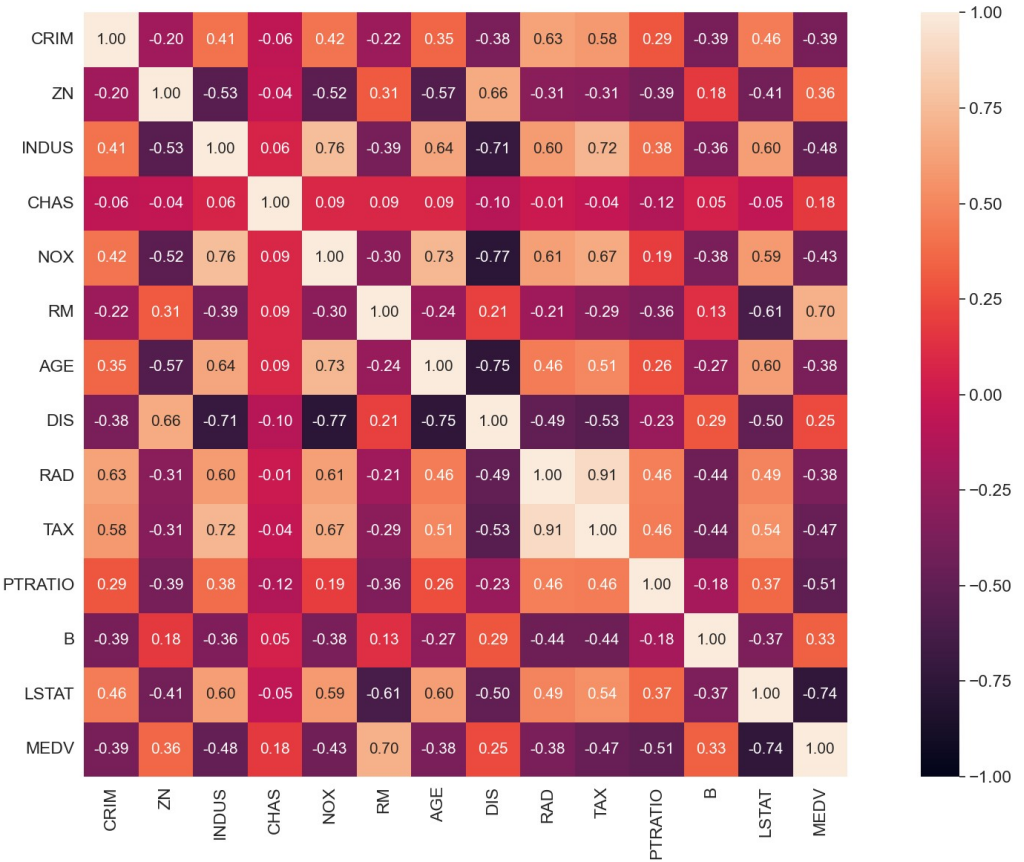
Increasing *bias* to decrease *variance* to increase *generalisability*

MODELS ASSUME INDEPENDENCE, BUT IT LOOK LIKE THERE IS NO INDEPENDENCE,
WHEN THERE IS OFF DIAGONAL VARIANCE (WE WANT IT TO BE 0, THE COVARIANCE)

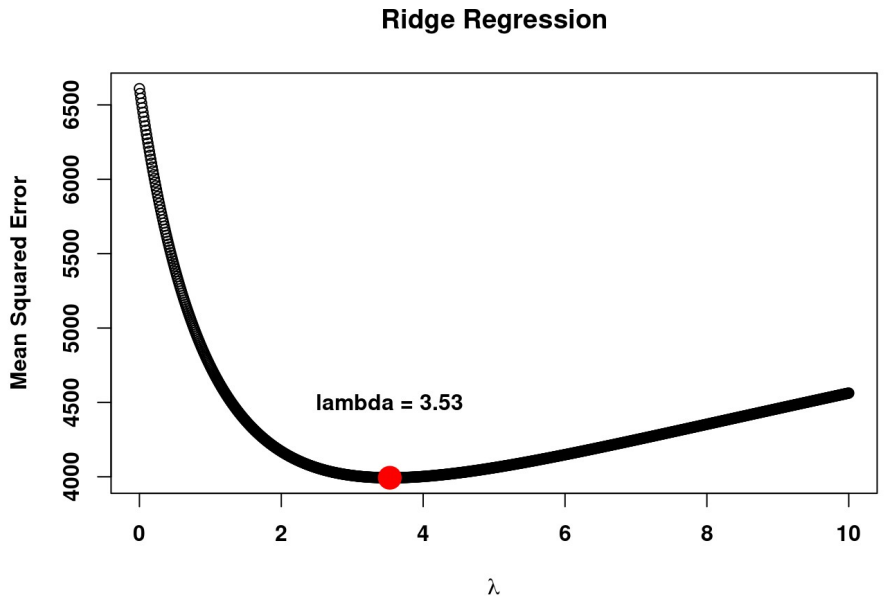
Off-diagonal variance

```
import numpy as np

X # feature matrix
cov features = np.cov(X.T)
```

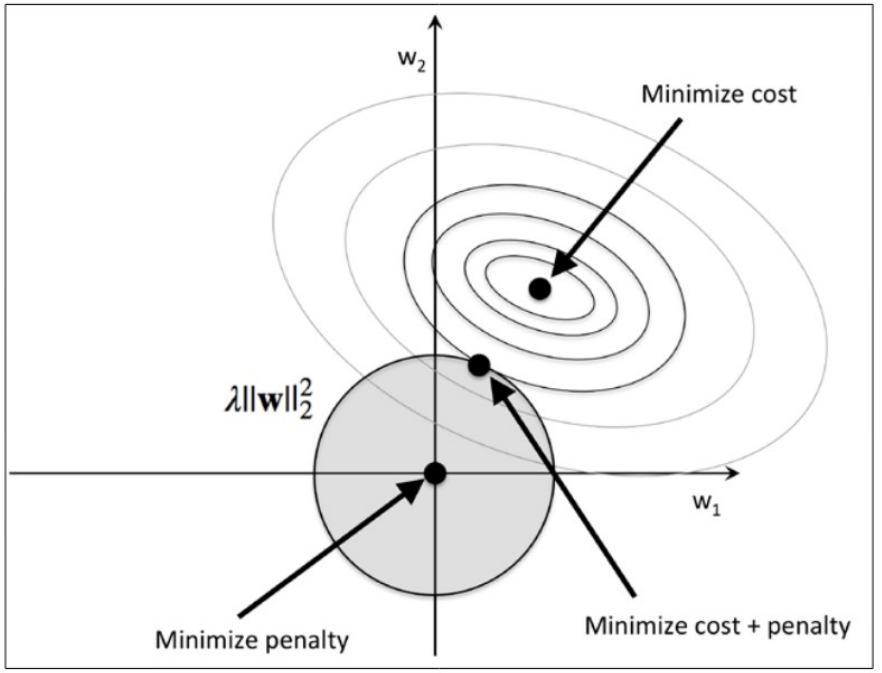




Decreasing variance through regularisation




```
from sklearn.linear_model import Ridge
alpha_range = np.arange(0.1, 100, 0.1)
mse_list = [None] * len(alpha_range) # initialise with
for alpha_index, alpha in enumerate(alpha_range):
    RR = Ridge(alpha=alpha)
    mse_list[alpha_index] = np.mean(cross_validate(RR, X_std, y_std, 11))
```



(p. 114: Raschka, 2015)

How gradient descent can be used to maximise likelihood in regression

A general formulation : $\Delta w_j = \eta \sum_i^n \left(y^{(i)} - \phi\left(z^{(i)} \right) \right) x_j^{(i)}$

Δw_j: updating of weight for parameter j

η: learning rate

y: target label

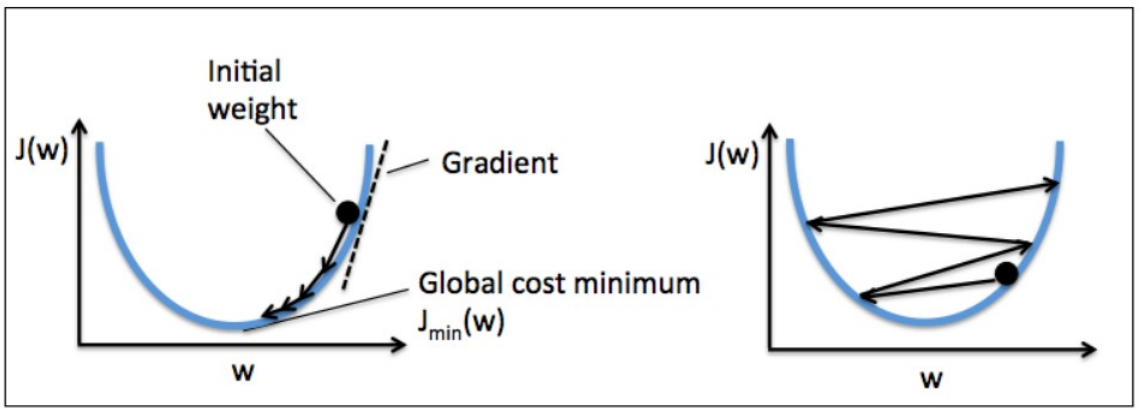
ϕ: link function

$\phi(z)$: predicted value

x: feature variables (predictor)

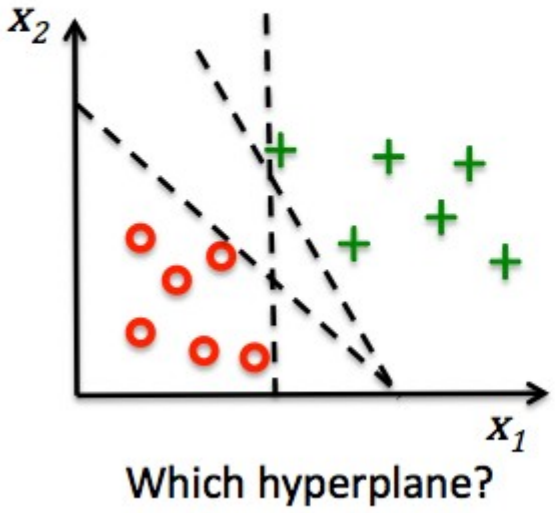
```python
class LinearRegressionGD(object):

    def __init__(self, eta=0.001, n_iter=20):
        self.eta = eta
        self.n_iter = n_iter

    def fit(self, X, y):
        self.w_ = np.zeros(1 + X.shape[1])
        self.cost_ = []

        for i in range(self.n_iter):
            output = self.net_input(X)
            errors = (y - output)
            self.w_[1:] += self.eta * X.T.dot(errors)
            self.w_[0] += self.eta * errors.sum()
            cost = (errors**2).sum() / 2.0
            self.cost_.append(cost)
        return self

    def net_input(self, X):
        return np.dot(X, self.w_[1:]) + self.w_[0]

    def predict(self, X):
        return self.net_input(X)
```

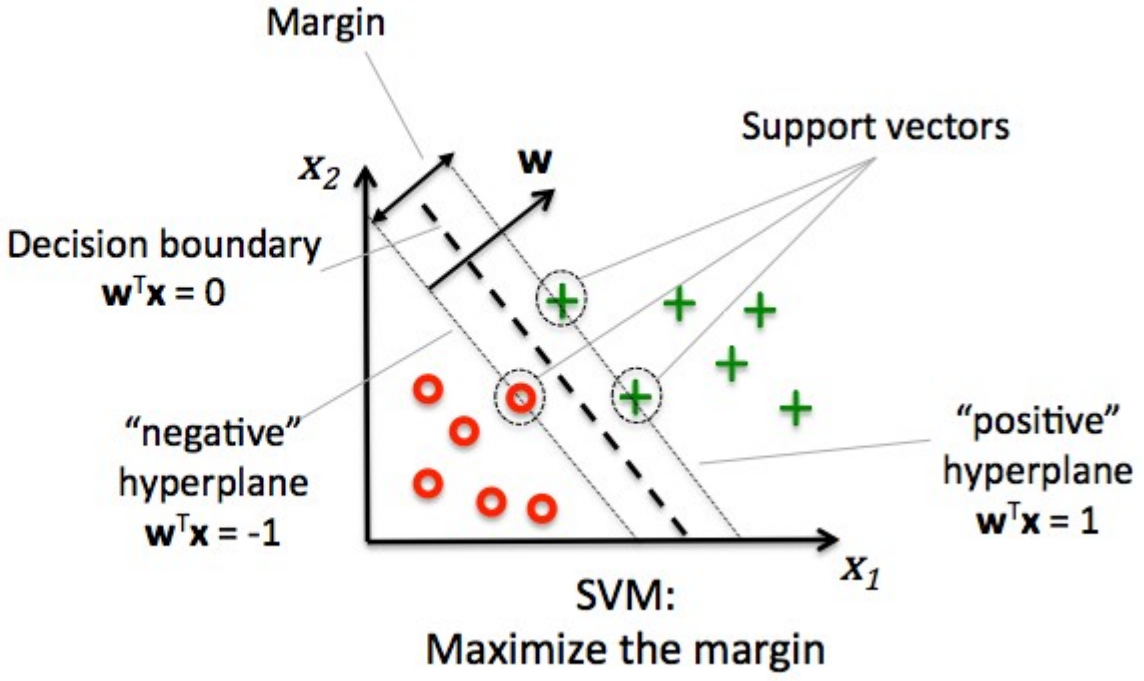
(p. 40: Raschka, 2015)



Introduced to (non-linear) Support Vector Machines

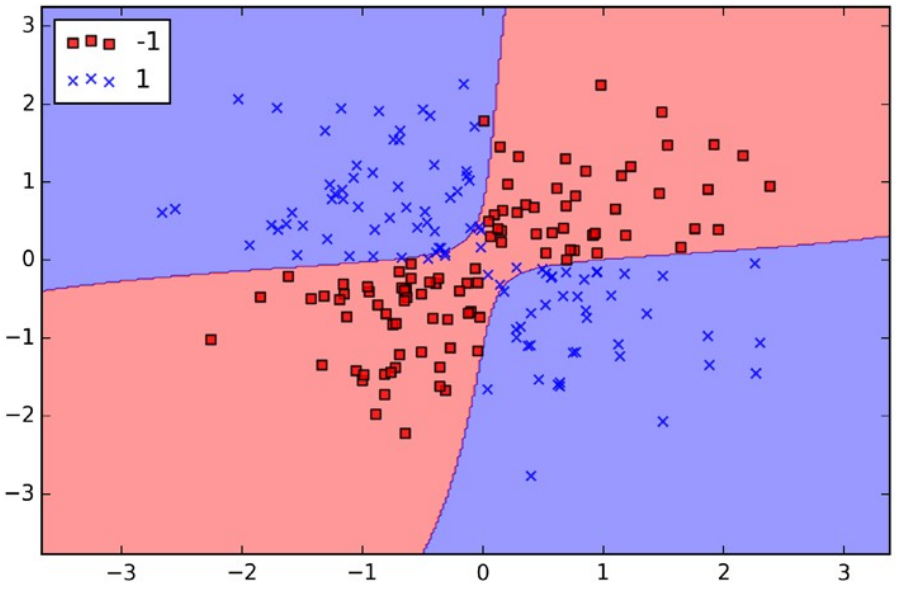
Support Vector Machine





(p. 69: Raschka, 2015)

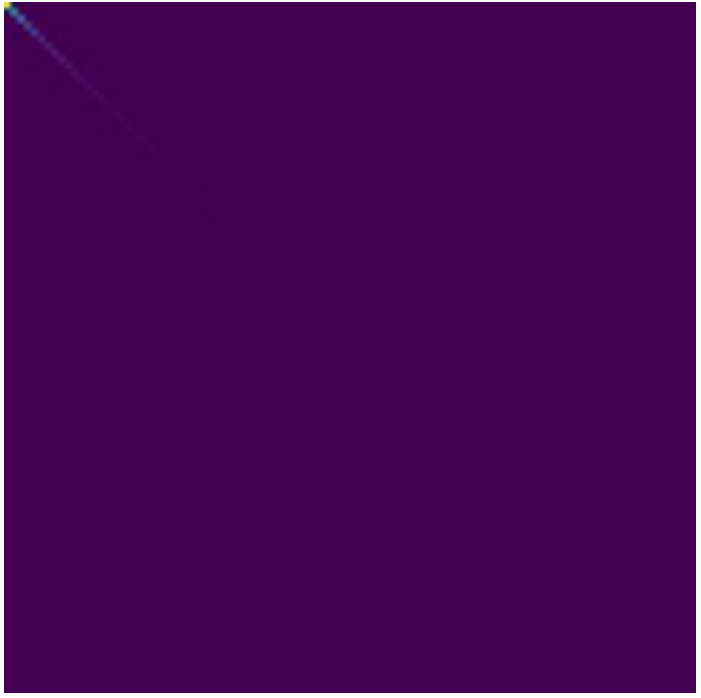
Support Vector Machine

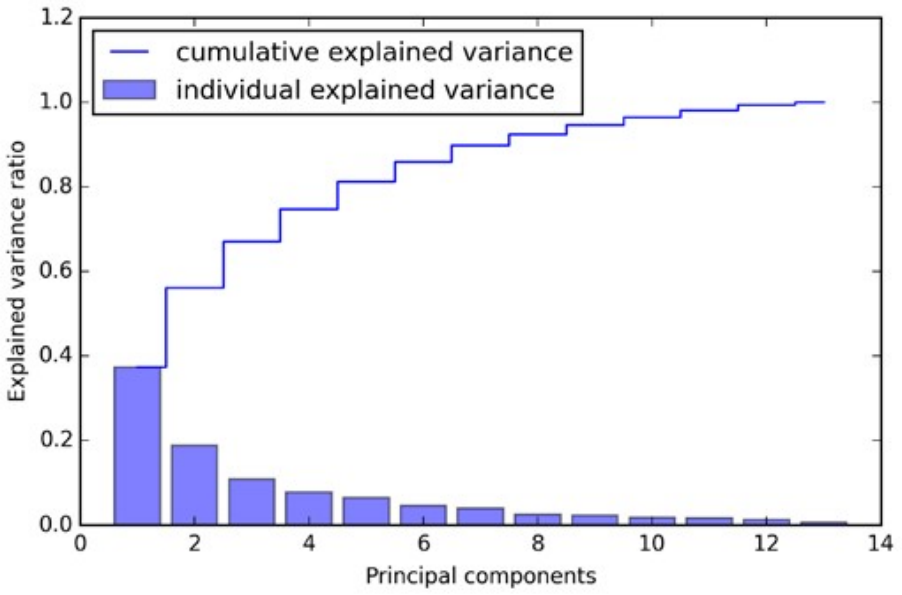


(p. 78: Raschka, 2015)

De-correlation and dimensionality reduction

Principal component analysis

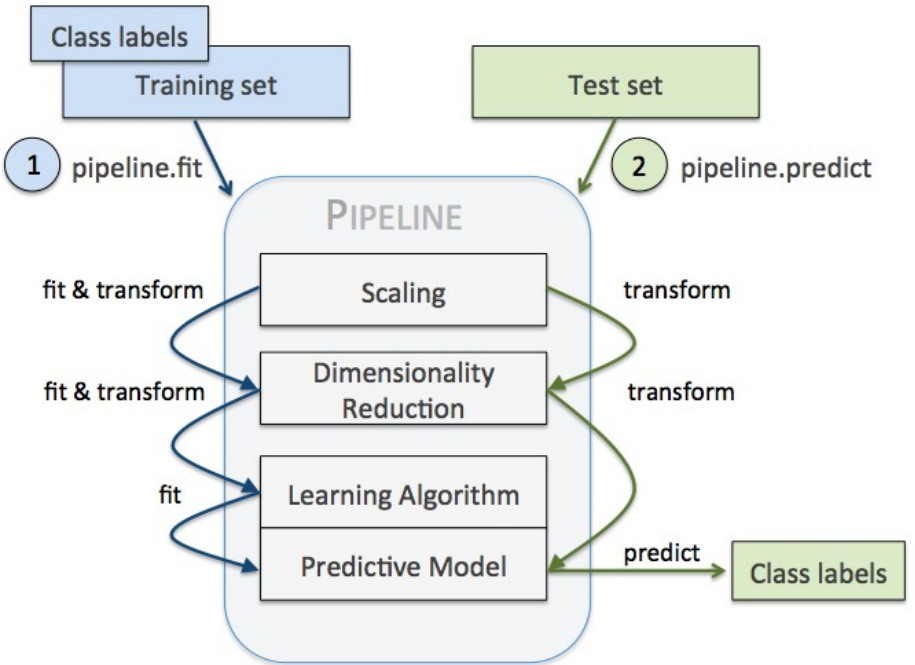





```
eigenvalues, eigevector = np.linalg.eig(cov features)
```

(p. 132: Raschka, 2015)

Machine learning pipelines

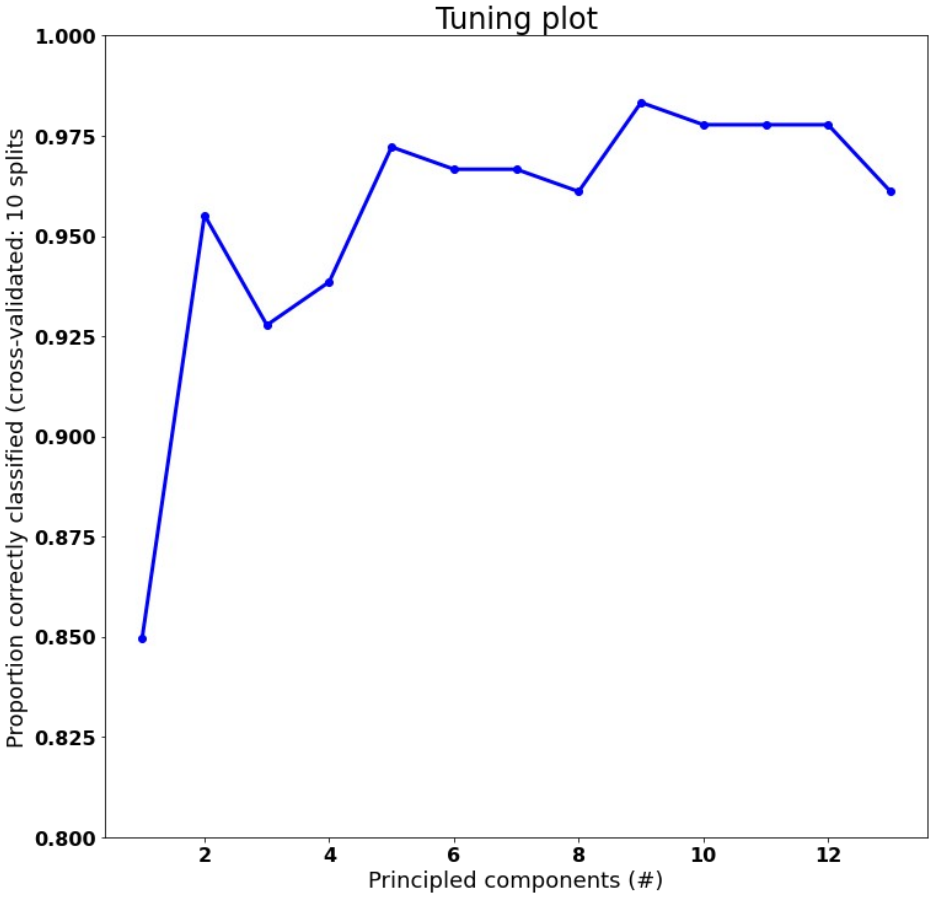


(p. 172: Raschka, 2015)

Cross-validation and tuning parameters

Also less of a problem with *K*-fold cross-validation

(p. 176: Raschka, 2015)



MACHINE LEARNING
You have learnt to increase the generalisability of models by introducing bias (e.g. regularisation) and reducing dimensionality (e.g. PCA and L1 regularisation)

Did we learn what we should?

Academic regulations – objectives

Knowledge:

- demonstrate understanding of statistical techniques relying on the Generalised Linear Model

- demonstrate understanding of hierarchical modeling methods

- demonstrate understanding of basic machine learning concepts.

Academic regulations – objectives

Skills:

- build and evaluate models of hierarchically structured data

- integrate machine learning procedures in data analysis

- communicate analysis processes, results and interpretation.

Academic regulations – objectives

Competences:

- independently decide on data analysis methods, given a data set and a research question

- justify decision making when pre-processing messy data for data analysis.

Thank you!

- for joining me on this ...
  - frustrating
  - challenging
  - time-demanding
- but hopefully also ...
  - rewarding
  - informative
  - challenging (!)

Journey towards knowledge, skill and competence

The only inspirational picture you got in a whole course: https://www.diygenius.com/self-education-will-make-you-a-fortune/

References

- Gelman, A., Hill, J., 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.
- Raschka, S., 2015. Python Machine Learning. Packt Publishing Ltd.
- Yarkoni, T., Westfall, J., 2017. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. Perspect Psychol Sci 12, 1100–1122. https://doi.org/10.1177/1745691617693393