Probabilistic Inference and Learning
Flipped Classroom - Week 1
Meeting starts 10:15 sharp. Please hold

Philipp Hennig

21 April 2020

Eberhard Karls
Universität Tübingen

Faculty of Science
Department of Computer Science
Chair for the Methods of Machine Learning

Goals for this Meeting

Flipped Classrooms every Tuesday, 10ct to 12

- First “symmetric” meeting
- Administrative Details
- Relationship to *Statistical ML*
- Some notes on conditional independence
- Your Questions

These slides will be on Ilias afterward.

Sources of Information

Let's stay connected

- Ilias for slides, exercise sheets, forum, communication
- youtube for videos (see below)

Don't trust Alma for anything at the moment!
For the ML MSc, this course is part of *Foundations of ML* and compulsory.

Lectures

Admin Slide 1

- lectures are available on youtube channel *Tübingen Machine Learning*

  `https://www.youtube.com/channel/UCupmCsCA5CFXmm31PkUhEbA`

- *two lectures per week*, available on **Friday, 6am**
- flipped classroom on Tuesdays, 10ct-12noon
  - participation is not compulsory
  - will often feature “nice to know” / “fun” content that is not required
  - use the opportunity to **ask your questions**!

Exercises

Admin Slide 2

- new exercise sheet every week, available Monday morning
- plenary tutorial session every Monday at 10ct-12noon
- exercises are **due every Monday at 10am sharp**

Each exercise sheet contains

- a simple "EXAMple" that can be answered more or less direclty from the course
- a **theory** question (math) that requires a bit of thinking
- a **practice** question (code), often a project extending across 1–5 sheets

Each exercise on your sheet will be graded as either *sufficient* or *insufficient*. To be marked *sufficient*, you must have made a serious effort to solve the task and have solved at least a nontrivial part of the task.

The Team

Each tutorial will be presented by the author of the coding exercise

Philipp Hennig
Methods of ML

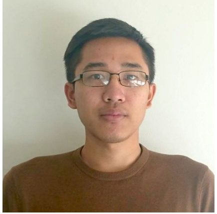
Filip de Roos
Gaussian Inference

Alexandra Gessner
Sampling

Julia Grosse
Sampling

Jonathan Wenger
Probabilistic Programming

Marius Hobbhahn
Bayesian Deep Learning



Agustinus Kristiadi
Approximate Inference

Nicholas Krämer
Approximate Inference

How to Submit your Exercises

please follow precisely

- submit on Ilias in folder labeled “upload your exercise sheets here”
- **submit a single pdf file** (for jupyter notebooks, export to pdf)
- name your file

  `XX_yoursurname_matrikelnummer.pdf`

  where xx is the number of the sheet. E.g.

  `01_Hennig_1033931.pdf`
- **up to two people can submit as a group, if you want**. Then name your sheet

  `XX_firstname_firstmatrikel_secondname_secondmatrikel.pdf`

Exam

the slide you've been waiting for

- Currently scheduled for 23 July 2020, 14ct-16:00
- Second exam scheduled for 05 October 2020, 10ct-12:00
- If you fail the first exam, you can take the second one
- If you do not participate in the first exam, you only have the second exam
- There are 12 exercise sheets with three exercises each
- To be **admitted to the exam** you must have scored “sufficient” on **at least 12 exercises** (i.e. 1/3 of the total)
- The exam will consist of a total of 100 points.
- **Every *coding* exercise scored as *sufficient* adds one (1) bonus point to your exam score** (i.e. up to 12 extra points, if you sufficiently solve all the coding exercises)
- You must pass the exam without (i.e. not counting) bonus points.
- Do not post solutions on the forum or elsewhere! Posting solutions will be counted as plagiarism and result in your entire sheet being graded as “insufficient.”

Questions about ADMIN matters?

Relationship to *Statistical Machine Learning*

Nb: These are separate courses

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

- Nb: The two courses are not timed or coordinated relative to each other. Don't expect us to make direct cross-links
- Nevertheless, their parallel delivery is by design
- There will be one or two dedicated lectures here to connect to specific parts of statistical machine learning (especially kernel methods)

Relationship to Other Approaches to Machine Learning

Why should you want to take this course?

Statistical Learning Theory	Probabilistic Learning
formulate a loss-function *ad hoc*, mapping data to predictions/decisions. Then carefully analyse to show that, under some external assumptions, this model has certain desirable properties.	Carefully formulate and critique a **generative model**. Inference is then uniquely determined by Bayes' theorem. No need to question or analyse the paradigm over and over again
mathematical **analysis** in the foreground (often in the asymptotic or large-number limit)	**numerical & computational** design in the foreground (the right model may be intractable)
statements about errors tend to focus on the **worst-case**	structured and extensive **quantification of uncertainty** by the posterior, often core motivation

Don’t be fight a pointless war!

Frequentists and Bayesians can get along.

The two frameworks do not contradict each other! Instead they complement each other:

Statistical Analysis can …

- reassure you that certain aspects of your prior don’t matter in the asymptotic limit, or that some assumptions are fundamentally ill-advised
- bound errors of estimators under quite general assumptions (e.g. just the σ-algebra, not the measure)

Probabilistic Analysis can …

- give intuition for how to choose *details* of a model and customize it to a specific task
- quantify uncertainty in finite time

In her first Lecture, Prof. von Luxburg explained that *inductive bias* is fundamentally required for machine learning.
In my first lecture, I mentioned that probabilistic reasoning can allow induction.

Some questions as the course progresses:

- Do all probabilistic methods have inductive bias?
- Is it possible to build models that “have no prior”?

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

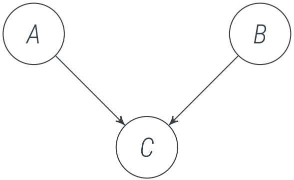
Definition (conditional independence)

Two variables A and B are ***conditionally independent*** *given variable C, if and only if their conditional distribution factorizes,*

$$P(A, B|C) = P(A|C)\ P(B|C)$$

In that case we have $P(A|B, C) = P(A|C)$, i.e. in light of information C, B provides no (further) information about A. Notation: $A \perp\!\!\!\perp B \mid C$

Example: Two coins and a bell.



$A =$ coin 1 shows heads

$B =$ coin 2 shows heads

$C =$ bell rings if both coins show the same result

$A \perp\!\!\!\perp B$ and $A \perp\!\!\!\perp C$ and $B \perp\!\!\!\perp C$, but $A \not\perp\!\!\!\perp B \mid C$ and $A \not\perp\!\!\!\perp C \mid B$ and $B \not\perp\!\!\!\perp C \mid A$.

The numerical values matter

Things change is the coins are not unbiased

Say $P(A=1) = 1/2$ $\qquad P(B=1) = 3/4$

Joint: $P(A,B,C)$:

$$C=0: \quad \begin{array}{cc} & B \\ & \begin{array}{cc} 0 & 1 \end{array} \\ A \begin{array}{c} 0 \\ 1 \end{array} & \left|\begin{array}{cc} 0 & 3/8 \\ 1/8 & 0 \end{array}\right| \end{array} \qquad C=1: \quad \begin{array}{cc} & B \\ & \begin{array}{cc} 0 & 1 \end{array} \\ A \begin{array}{c} 0 \\ 1 \end{array} & \left|\begin{array}{cc} 1/8 & 0 \\ 0 & 3/8 \end{array}\right| \end{array} \Rightarrow$$

Marginal $P(A,B)$:

$$\begin{array}{cc} & B \\ & \begin{array}{cc} 0 & 1 \end{array} \\ A \begin{array}{c} 0 \\ 1 \end{array} & \left|\begin{array}{cc} 1/8 & 3/8 \\ 1/8 & 3/8 \end{array}\right| \end{array} = \underbrace{\left|\begin{array}{c} 1/2 \\ 1/2 \end{array}\right|}_{P(A)} \cdot \overbrace{\left|\begin{array}{cc} 1/4 & 3/4 \end{array}\right|}^{P(B)}$$

$\Rightarrow A \perp\!\!\!\perp B$

But $P(A,C)$:

$$\begin{array}{cc} & C \\ & \begin{array}{cc} 0 & 1 \end{array} \\ A \begin{array}{c} 0 \\ 1 \end{array} & \left|\begin{array}{cc} 3/8 & 1/8 \\ 1/8 & 3/8 \end{array}\right| \end{array} \Rightarrow A \not\perp\!\!\!\perp C$$

yet $P(B,C)$:

$$\begin{array}{cc} & C \\ & \begin{array}{cc} 0 & 1 \end{array} \\ B \begin{array}{c} 0 \\ 1 \end{array} & \left|\begin{array}{cc} 1/8 & 1/8 \\ 3/8 & 3/8 \end{array}\right| \end{array} \Rightarrow B \perp\!\!\!\perp C$$

To show (conditional) independence, you have to consider the entire joint probability table.

Your Questions!

Are you Ok?

Contacts

- for psychological problems: the Psychosoziale Beratungsstelle of the Studentenwerk (also in English!) `pbs-stuwe@sw-tuebingen-hohenheim.de`
- regarding your degree:
  - Prof. Dr. Matthias Hein (MSc ML) `Matthias.Hein@uni-tuebingen.de`
  - your degree coordinator or Prüfungssekretariat[1]
  - Prof. Dr. Kay Nieselt (Dean of Studies) `nieselt@informatik.uni-tuebingen.de`
- if your are feeling physically ill, **call** your general practitioner, or the Fieberambulanz on Festplatz (daily, 9-17:00. 07071 207-3600)

Make sure to keep personal contact with your fellow students!

[1] `http://www.informatik.uni-tuebingen.de/studium/ansprechpartner-und-organisation/pruefungssekretariate/`