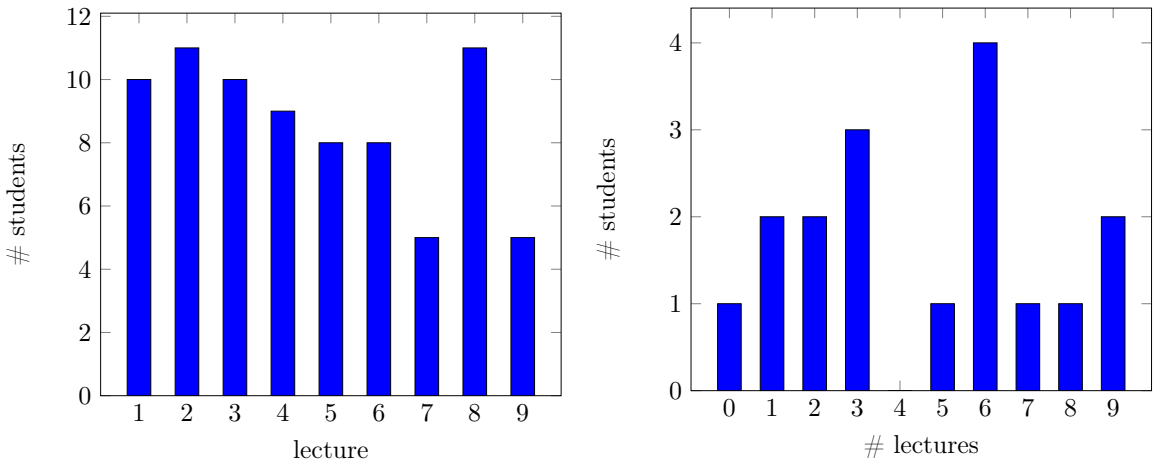
KTH Computer Science and Communication

# DT2118: Speech and Speaker Recognition
# Course Analysis VT2014

## 1 Course data

| | |
|---|---|
| Course name | Speech and Speaker Recognition |
| Course number | DT2118 |
| Credits | 7.5 |
| When the course took place | Period 4, VT2014 |
| Teachers (hours) | Giampiero Salvi (course responsible and lecturer, 18F, 2T, 4Ö, 8L, 4S)[1] |
| Number of registered students | 15 |
| Number of students attending | 11+2PhD students |
| Number of students completing | 9 |




Although attendance to the lectures was not mandatory, I kept a record of the students who participated to each lecture. The left plot above shows the number of students for each of the 9 lectures. The right plot shows a histogram of how many students attended a certain number of lectures. Only one student was never present, and only two students attended all lectures.

## 2 Course objectives

The course objective is to give students insights in the signal processing and statistical methods used for speech and speaker recognition. After attending the course, the students should be able to:

[1] F: lectures (föreläsningar), T: tutorial, Ö: exercises (övningar), L: lab, S: final seminar

1. discuss concepts related to speech signal processing and speech recognition
2. explain and implement simple parameter estimation methods for Gaussian models
3. explain and implement the main algorithms related to hidden Markov model training and decoding
4. use the software package KTH to build and evaluate a simple speech recogniser
5. carry out a small project related to speech or speaker recognition

# 3 How the course is designed to reach the objectives

The course gives a central space to the role of student activities as a means to learning. The 20 hours of lectures are meant to give the context and sufficient insights to the students in order to perform the different kinds of activities. Among these are: a set of computational exercises, a laboratory exercise and a final project.

The computational exercises require a mixture of theoretical and programming skills. The students are not only required to solve the exercises theoretically and numerically, but also to hand in the code they have written for the solution. Matlab is the preferred language, but the students are left free to choose their favourite programming language. A central part of this activity has been the discussion with the teacher after the solution has been handed in. Given the relatively low number of students, the course responsible was able to sit with each student and check their solutions in details, searching for eventual bugs in their code.

The laboratory is based on a software package for speech recognition (HTK) and on recordings made by the students in a previous lab (from DT2112). Students work in groups of about 2. Students that did not participate in DT2112 are required to go through the DT2112 lab before they start the DT2118 lab. The aim of this exercise is to compare the effects of different feature extraction methods on the recognition accuracy in a digit recognition task. As an optional task, the students are asked to perform Vocal Tract Length Normalisation and to discuss the results they obtain.

The final activity is the final project. I tried to keep the subject of the projects as open as possible for two reasons. The first is the choosing the right project on the base of the information given by the lectures and one's own skills is in my opinion a learning activity *per se*. The other reason is that this kind of project work is best performed when the students are highly motivated, and being able to choose the task is a strong prerequisite to motivation. Of course, the students need continuous input from the teachers in order to realise if their ideas are feasible with the limited time given by the course. The students are required to work in groups of about 2 students (but also single student projects have been allowed this year). The need to perform the experiments related to the task they have chosen and write a report. Each student is also supposed to review the work of another group. In the final meeting in the course, each group is required to present the project, and the reviewers are required to ask related questions. Also the rest of the students are stimulated to participate in the discussion. In order to help the reviewers prepare for the discussion, each groups has been asked to hand in a draft of the report already one week before the final meeting.

# 4 Course pedagogical development I

This year I was responsible for all the activities in the course. The material in the course is updated every year. The main improvement was the development of an introduction to basic

signal theory (signal processing). This is a fundamental precondition for the course, but, unfortunately, the majority of students attending Computer Science programmes do not have courses in Signal Processing. I therefore added and introduction to the basic concepts that are required to understand key aspects of this course. In future years, this introduction should be extended and probably the number of lectures should be accordingly increased.

The laboratory instructions have also been updated, and the scripts have been made compatible with the Ubuntu installations at CSC giving the students more freedom on where to preform the assignment. Also the task has been redesigned in order to speed up the computer processing and leave more time to the students to perform learning activities during the lab.

On the down side, because of being also course responsible for DD1371, I was not able this year to sit with each student giving feedback on the computational exercises. This showed in some of the comments in the questionnaire.

# 5 Contact with the students during the course

## 5.1 Students in this year's committee

This year no student representatives were elected for lack of availability of the students.

## 5.2 Results of the course meeting

There was no formal meeting during the course, but there was always close discussion with the students, also simplified by the low number of students.

# 6 Contact with the other teachers during the course

This point is not relevant being I the only teacher.

# 7 Questionnaire, the student's point of view

## 7.1 Period in which the questionnaire was active

After the last meeting of the course, until the final meeting with the course responsible (about two weeks)

## 7.2 Questions that were added to the standard

See the questionnaire

## 7.3 Statistics of answers

5/11 active students answered the questionnaire.

## 7.4 Changes compared to the last implementation

No modifications were made this year.

### 7.5 General impression

The general impression on the course is positive. Students appreciate especially the computational exercises that give insights in the details of the algorithms, and the fact that the teacher was always available for discussion (this is of course dependent on the number of students that enrolled this year).

Among the written comments there is some constructive criticism, but always combined with positive forced choice answers.

### 7.6 Relevant web links

## 8 Interpretation of the questionnaire by the course responsible

### 8.1 Positive views

- most students are happy about the course organization
- they find the course interesting
- all but one find the lectures good or very good pedagogically
- all students think the material was good or very good

### 8.2 Negative views

- Many students think the theoretical aspects should be dealt with in more details
- some students would like to have more feedback on the computational exercise
- 2 out of 3 students think the course should be less than 7.5 credits

### 8.3 Was the course relevant with respect to the objectives

Yes

### 8.4 View on prerequisites

The prerequisite of DT2112 should be removed as it is not necessary for this course. Some students were not aware of it and many did not attend DT2112 and completed DT2118 successfully.

### 8.5 View on forms of teaching

### 8.6 View on literature and course material

### 8.7 Opinion on exam

### 8.8 Especially interesting comments

- “I really enjoyed the exercises. My only request is for some debugging information to be provided, e.g. "if you have implemented algorithm X correctly, you should get the following result when you feed it the following data...(example data and correct result provided).”
- “It is not really required to understand anything to complete the labs...”
- “Perhaps more on deep learning neural networks.”

# 9 Opinions from the other teachers after the end of the course

## 9.1 What worked well

## 9.2 What worked less well

## 9.3 Suggestions for improvements

# 10 Results of the course commission meeting after the exam

## 10.1 Summary from the students

## 10.2 Suggestions for changes

## 10.3 Link to the meeting protocol

# 11 Summary of the course responsible report

## 11.1 General impression

The general impression of the course is positive.

## 11.2 Positive points of view

## 11.3 Negative points of view

## 11.4 Opinion on prerequisites

The prerequisite are satisfactory. Some details will be adjusted.

## 11.5 Opinion on forms of teaching

The forms of teaching seem to be satisfactory (see also previous points)

## 11.6 Opinion on literature/course material

The literature seems to be satisfactory.

## 11.7 Opinion on examination

The examination seems to have worked fine.

# 12 Course pedagogical development II

## 12.1 How the changes to this course work

The changes to the previous years seem to have worked well.

## 12.2 Changes that should be done for the next time

- increase the introduction to Signal Processing (adding one lecture to the course)
- improve the labs by adding steps that require implementation of ASR methods.
- improve the computational exercises by giving test data to debug the algorithms.

- possibly, add topics on deep learning in the lectures.
- add a lecture that puts together all the different parts learned during the course.

# DT2118 Speech and Speaker Recognition

## Resultat av kursutvärdering

---

**25 questions to be answered and a few comments to be given.**

---

1. Do you think the course is easy or difficult?

    1. 0% (0 st) Very easy.
    2. 40% (2 st) Easy.
    3. 40% (2 st) Average.
    4. 20% (1 st) Rather difficult.
    5. 0% (0 st) Very difficult.

---

2. Did you get a clear idea of the course objective at the course start?

    1. 100% (5 st) Yes.
    2. 0% (0 st) Hesitant.
    3. 0% (0 st) No.

---

3. Do you think the course is interesting?

    1. 20% (1 st) Yes, very.
    2. 60% (3 st) Yes.
    3. 20% (1 st) Neutral.
    4. 0% (0 st) Not very.
    5. 0% (0 st) No.

---

4. How did you get to know about this course?

    1. 0% (0 st) Friend.
    2. 100% (5 st) KTH web pages.
    3. 0% (0 st) From DT2112.
    4. 0% (0 st) Other (please specify).

    Comments:

---

5. The requirement on previous courses is the Speech Technology course DT2112 or equivalent and experience with unix. Do you regard your level as sufficient at the time of the course start?

    1. 80% (4 st) Yes.
    2. 20% (1 st) Hesitant.
    3. 0% (0 st) No.

    Comments:

    *The required experience in Unix is so basic, that it shouldn't even be listed for most KTH students - it is safe to assume that they have it.*
    *As far as DT2112: I did not take this class previously, but I had no problem catching up.*
    ---
    *I didn't take DT2112 but have seen some of the speech stuff from a course on an exchange semester.*

---

6. What is your opinion on the course book "Spoken Language Processing"?

    1. 0% (0 st) Very good.
    2. 20% (1 st) Good.
    3. 0% (0 st) Acceptable.
    4. 0% (0 st) Not so good.
    5. 0% (0 st) Bad.
    6. 80% (4 st) Did not use it.

    Comments:

    *Did not use it yet.*

---

7. What is your opinion on the extra material (papers, handouts, etc)?

    1. 40% (2 st) Very good.
    2. 60% (3 st) Good.
    3. 0% (0 st) Acceptable.
    4. 0% (0 st) Not so good.
    5. 0% (0 st) Bad.
    6. 0% (0 st) Did not use it.

    Comments:

    *giving out slide handouts at the start of the lecture are very handy for taking notes.*

---

8. Was there a good balance between lectures and practical activities?

    1. 20% (1 st) Too many lectures.
    2. 80% (4 st) Good balance.
    3. 0% (0 st) Too many practical activities.

    Comments:

    *I have often benefited from smaller but more frequent labs, then one or two big labs. On the other hand spending more time on a project instead of smaller assignments would also be good in some sense. I believe this should be governed by the goal of the course. If at the end of the course students should be able to implement the Viterbi then it should be part of the exercises. (I am yet to do the two labs.)*

---

9. What do you think of the lectures in a pedagogical way? (Are the concepts well described? Do the teachers speak and write clearly)

    1. 40% (2 st) Very good.
    2. 40% (2 st) Good.
    3. 0% (0 st) Acceptable.
    4. 20% (1 st) Not so good.
    5. 0% (0 st) Bad.

Comments:

*The lectures are not so clear. It is often hard to get a good idea of the algorithms and how various parts work together... It may lack some details.*

---

10. Do you think the theoretical aspects should be expanded more in details?

    1. 20% (1 st) Fully agree.
    2. 20% (1 st) Agree.
    3. 60% (3 st) Do not have an opinion.
    4. 0% (0 st) Disagree.
    5. 0% (0 st) Fully disagree.

    Comments:

    *I think a few actual simple examples would be extremely helpful when explaining the more complex algorithms.*

---

11. How was your work load distributed during the course?

    1. 0% (0 st) Mainly at the beginning.
    2. 40% (2 st) Evenly during the course.
    3. 60% (3 st) Mainly at the end.

    Comments:

---

12. The assessment in this course is based on a Fail/Pass grading scale. Do you think we should assign more specific grades? For example on a 7 level scale?

    1. 0% (0 st) Yes.
    2. 0% (0 st) Hesitant.
    3. 100% (5 st) No.

    Comments:

    *I think the class is great the way it is.*

---

13. What is your opinion on the computational exercises?

    1. 20% (1 st) Very good.
    2. 20% (1 st) Good.
    3. 0% (0 st) Acceptable.
    4. 20% (1 st) Not so good.
    5. 0% (0 st) Bad.
    6. 40% (2 st) No opinion.

    Comments:

    *I really enjoyed the exercises. My only request is for some debugging information to be provided, e.g. "if you have implemented algorithm X correctly, you should get the following result when you feed it the following data...(example data and correct result provided).*
    ---
    *I am yet to finish them.*
    ---
    *It is only Machine Learning and would benefit from being more ASR related.*

---

14. What is your opinion on the practical exercise (lab)?

    1. 0% (0 st) Very good.
    2. 40% (2 st) Good.
    3. 0% (0 st) Acceptable.
    4. 0% (0 st) Not so good.
    5. 20% (1 st) Bad.
    6. 40% (2 st) No opinion.

    Comment:

    *I am yet to finish them.*
    ---
    *It is not really required to understand anything to complete the labs...*

---

15. How long time did you spend preparing the practical exercise (lab)?

    1. 40% (2 st) Less than 6 hours.
    2. 0% (0 st) 6-12 hours.
    3. 20% (1 st) 12-24 hours.
    4. 20% (1 st) More than 24 hours.

    Comment:

    *I am yet to finish them.*

---

16. How long time did you spend on writing the term paper and preparing the presentation:

    1. 0% (0 st) Less than one day.
    2. 20% (1 st) 2-3 days.
    3. 40% (2 st) 4-6 days.
    4. 40% (2 st) 7 days or more.

    Comments:

---

17. How many other courses did you follow in parallel to this one (period 4)?

    1. 20% (1 st) None.
    2. 20% (1 st) One.
    3. 20% (1 st) Two.
    4. 40% (2 st) Three.
    5. 0% (0 st) Four or more.

---

18. How large proportion of your studying time in period 4 did you spend on this course?

    1. 20% (1 st) Less than 15%.
    2. 40% (2 st) 15-30%.
    3. 20% (1 st) 30-50%.
    4. 20% (1 st) 50-70%.
    5. 0% (0 st) More than 70%.

---

19. The course is 7.5 hp. How do you regard that in comparison with other courses?

    1. 60% (3 st) 7.5 hp is appropriate.
    2. 0% (0 st) Should be more than 7.5 hp.
    3. 40% (2 st) Should be less than 7.5 hp.

---

20. What is your opinion on the administration of the course and exchange of information between teachers and students?

    1. 80% (4 st) Very good.
    2. 0% (0 st) Good.
    3. 0% (0 st) Acceptable.
    4. 20% (1 st) Not so good.
    5. 0% (0 st) Bad.

    Comments:

    *No feedback on the computational exercises.*

---

21. Were the requirements at every stage of the course always clear to you?

    1. 60% (3 st) Yes.
    2. 40% (2 st) Hesitant.
    3. 0% (0 st) No.

    Comments:

---

22. Do you feel that you have been discriminated in this course due to gender, sexuality, ethniticity or disability?

    1. 0% (0 st) Yes.
    2. 0% (0 st) Hesitant.
    3. 100% (5 st) No.

    If yes, in which way?

---

23. How do you regard the course from a gender perspective (e.g. with respect to course book, teachers, etc.)?

    *No issues there.*
    ---
    *No opinion*
    ---
    *This thought never occured to me.*

---

24. Suggestion for course improvements:

    *One suggestion: provide sample input and answers to the programming exercises (some debugging info), so we can know whether or not we implemented the algorithms correctly.*
    ---
    *Include an executive summary and/or a list of keywords or key concepts for each lecture*
    ---
    *Follow up on the computational exercises.*
    ---
    *Perhaps more on deep learning neural networks.*
    ---
    *The number of lectures could be increased to focus more time on certain topics.*
    ---
    *Explain more how all components relate to each other in an ASR system.*

---

25. Further comments:

---

***giampi@kth.se***