



KTH Computer Science  
and Communication

# DT2118: Speech and Speaker Recognition

## Course Analysis VT2011

### 1 Course data

Course name	Speech and Speaker Recognition
Course number	DT2118
Credits	7.5
When the course took place	Period 4, VT 2011
Teachers (hours)	Giampiero Salvi (course responsible and lecturer, 4F, 2T, 4Ö, 4L, 2S) <sup>1</sup> Mats Blomberg (lecturer, 6F, 4L, 2S), Kjell Elenius (lecturer, 2F), Kornel Laskowski (lecturer, 2F)
Number of registered students	15(+1) <sup>2</sup>
Number of students attending	11
Number of students completing	7 <sup>3</sup>

### 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. 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 seven lectures are meant to give the context and sufficient insights to the students in order to

<sup>1</sup>F: lectures (föreläsningar), T: tutorial, Ö: exercises (övningar), L: lab, S: final seminar

<sup>2</sup>One student enrolled from Patentverket and never officially registered

<sup>3</sup>Four students will complete one of the assignments in Autumn 2011

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

For this year implementation of the course, a new introductory lecture has been added in order to give a general picture of the subject, especially for the students that did not attend DT2112. A second lecture has been redesigned to focus more on theoretical aspects of machine learning that are relevant to the course.

The computational exercises have been updated with new theoretical questions, and an exercise package has been developed including a number of Matlab prototype function to guide the implementation of the functions that are necessary to solve the exercises.

The laboratory instructions have also been updated.

## 5 Contact with the students during the course

### 5.1 Students in this year's committee

Niklas Vanhainen and Urban Thunberg

## 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

The communication with the other teachers was always very smooth, being us all colleagues from the same research group.

## 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

85% of the students that completed all the course requirements, and 55% of the total.

### 7.4 Changes compared to the last implementation

One question was removed compared to the last implementation, because irrelevant: "This year (2010), the lectures were replaced by self-studies in the course book, followed by discussions with the teacher. How do you regard this format?"

Eight questions were added compared to the last implementation:

1. How did you get to know about this course?
2. What is your opinion on the extra material (papers, handouts, etc)?
3. Was there a good balance between lectures and practical activities?
4. What do you think of the lectures in a pedagogical way? (Are the concepts well described? Do the teachers speak and write clearly)
5. How was your work load distributed during the course?
6. 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?
7. What is your opinion on the administration of the course and exchange of information between teachers and students?
8. Were the requirements at every stage of the course always clear to you?

## 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 teachers were always available for discussion (this is of course dependent on the number of students that enrolled this year).

## 7.6 Relevant web links

# 8 Interpretation of the questionnaire by the course responsible

## 8.1 Positive views

- All students but one are happy about the balance between standard lectures and practical activities.
- Students are very happy about the computational exercises. Also the choice of the course responsible to sit with them individually and check for mistakes and bugs was very much appreciated.
- Students liked in general the availability and engagement of the teachers
- Students are happy about the organisation of the course

## 8.2 Negative views

- One student complained that there should be more standard lectures in the course (but not less practical activities)
- The course book was not used by half of the students who filled the questionnaire
- Some students found the lab description not very clear
- Half of the students state that their work load was mainly concentrated at the end of the course.

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

Yes

## 8.4 View on prerequisites

They are appropriate, besides a detail about the use of emacs that was a heritage of older implementations of the course and will be removed in the future.

## 8.5 View on forms of teaching

One students think there should be more lectures

## 8.6 View on literature and course material

Many students did not use it and found the information they needed on the web

## 8.7 Opinion on exam

Most students are positive about the project presentations

## 8.8 Especially interesting comments

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

## 9.1 What worked well

Positive interest and response from the students

## 9.2 What worked less well

- The lecture schedule did not fit all students.
- One (or a few?) lecture(s) had to be given in a quite small lecture room

## 9.3 Suggestions for improvements

- Graded examination
- Practical exercises on other components than acoustic modeling

# 10 Results of the course commission meeting after the exam

At the meeting we went through the results of the questionnaire and the student representatives helped the course responsible to interpret them and added some suggestions and viewpoints.

## 10.1 Summary from the students

## 10.2 Suggestions for changes

- More example of scientific papers
- Make the course book as optional
- give more specific information about the parts of the book that are essential
- more lectures on search algorithms
- anticipate the deadline for the project report draft to force students to start earlier
- use a 7 level grading scale

## 10.3 Link to the meeting protocol

Later in this PDF

# 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**

There seems to be a good balance between lectures and practical assignments. The availability of the teachers seems to be highly appreciated by the students. The organisation of the different steps in the course seems to have worked well. The students were very motivated, they attended the lectures regularly (although it was not mandatory) and participated actively in the discussions. They respected all the deadlines, and if not possible they discussed in time the possibility of postponing a certain assignment to another termin.

## **11.3 Negative points of view**

There is room for at least one more lecture. Some assignments could be anticipated in order to distribute better the work load. Many students chose the literature study instead of the project because of lack of time.

## **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 course book was not extensively used by the students. The extra material in form of articles was more appreciated.

## **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. Especially the decision to check the exercise solutions individually with the students.

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

- one more lecture will be added
- the deadline for the draft report will be anticipated in order to force the students to start early on the projects
- some ready made material for the project will be made available in order to simplify the work
- the course results might be graded on a seven grade scale
- the lab description will be revised and clarified

- the lab scripts will be modified to make sure they can be run on standard Linux computers available at CSC
- the lectures by Kjell Elenius and Kornel Laskowski will be given by Giampiero Salvi (because both teachers do not work at our lab any longer)

# DT2118 Speech and Speaker Recognition

## Results of Course Evaluation

---

**24 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. 17% (1 st) Easy.
  3. 83% (5 st) Average.
  4. 0% (0 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. 50% (3 st) Yes.
  2. 33% (2 st) Hesitant.
  3. 17% (1 st) No.
- 

3. Do you think the course is interesting?

1. 17% (1 st) Yes, very.
  2. 50% (3 st) Yes.
  3. 33% (2 st) Neutral.
  4. 0% (0 st) Not very.
  5. 0% (0 st) No.
- 

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

1. 17% (1 st) Friend.
  2. 67% (4 st) KTH web pages.
  3. 17% (1 st) From DT2112.
  4. 0% (0 st) Other (please specify).
- 

Comments:

*From KTH web Pages, from a Friend, and from DT2112*

---

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



start?

1. 67% (4 st) Yes.
2. 33% (2 st) Hesitant.
3. 0% (0 st) No.

Comments:

*Never quite understood what good the emacs requirement that pops up every now and then is though ;)*

---

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

1. 0% (0 st) Very good.
2. 50% (3 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 50% (3 st) Did not use it.

Comments:

*Mostly found the things I was looking for regarding the basics at least.*

---

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

1. 0% (0 st) Very good.
2. 67% (4 st) Good.
3. 0% (0 st) Acceptable.
4. 17% (1 st) Not so good.
5. 0% (0 st) Bad.
6. 17% (1 st) Did not use it.

Comments:

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

1. 0% (0 st) Too many lectures.
2. 83% (5 st) Good balance.
3. 0% (0 st) Too many practical activities.

Comments:

*Too few lectures. Crammed up.*

---

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. 17% (1 st) Very good.
2. 33% (2 st) Good.
3. 17% (1 st) Acceptable.
4. 17% (1 st) Not so good.
5. 0% (0 st) Bad.

Comments:

---

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

1. 0% (0 st) Mainly at the beginning.
2. 50% (3 st) Evenly during the course.
3. 50% (3 st) Mainly at the end.

Comments:

---

11. 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. 33% (2 st) Yes.
2. 33% (2 st) Hesitant.
3. 33% (2 st) No.

Comments:

---

12. What is your opinion on the computational exercises?

1. 67% (4 st) Very good.
2. 33% (2 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 0% (0 st) No opinion.

Comments:

*It helped me understand all the concepts much much better as I had to implement them from scratch. Giampi had a lot of patience to teach us everything and debug with us.*

---

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

1. 17% (1 st) Very good.
2. 67% (4 st) Good.
3. 17% (1 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 0% (0 st) No opinion.

Comment:

*More relevant explanation was required before the start of the actual lab*

---

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

1. 50% (3 st) Less than 6 hours.
2. 17% (1 st) 6-12 hours.
3. 17% (1 st) 12-24 hours.
4. 17% (1 st) More than 24 hours.

Comment:

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

1. 0% (0 st) Less than one day.
2. 33% (2 st) 2-3 days.
3. 17% (1 st) 4-6 days.
4. 33% (2 st) 7 days or more.

Comments:

*Not so much the writing as the project*

---

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

1. 0% (0 st) None.
  2. 17% (1 st) One.
  3. 67% (4 st) Two.
  4. 17% (1 st) Three.
  5. 0% (0 st) Four or more.
- 

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

1. 0% (0 st) Less than 15%.
2. 50% (3 st) 15-30%.
3. 33% (2 st) 30-50%.

4. 0% (0 st) 50-70%.
  5. 17% (1 st) More than 70%.
- 

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

1. 100% (6 st) 7.5 hp is appropriate.
  2. 0% (0 st) Should be more than 7.5 hp.
  3. 0% (0 st) Should be less than 7.5 hp.
- 

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

1. 83% (5 st) Very good.
2. 17% (1 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.

Comments:

*Giampi and Mats were easily accessible and answered any questions thoroughly and well.*

---

*Matts and Giampi are really very helpful understanding teachers, with a lot of patience. They were eager to teach us and always very accommodating and approachable.*

---

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

1. 67% (4 st) Yes.
2. 33% (2 st) Hesitant.
3. 0% (0 st) No.

Comments:

---

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

1. 0% (0 st) Yes.
2. 0% (0 st) Hesitant.
3. 100% (6 st) No.

If yes, in which way?

---

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

*Nothing to comment.*

---

23. Suggestion for course improvements:

*More lectures and covering the speech concepts much more in detail.*

---

24. Further comments:

*The best course I have taken during my years at KTH bar none. Don't change a thing! :)*

*I really liked the open nature of the project/term paper part of the course, I learned a lot more from it than I would have from any written exam. The teachers could answer any questions regarding the subject and were easily accessible and easy to talk to. The small class size probably didn't hurt either.*

---

[giampi@kth.se](mailto:giampi@kth.se)

Denna sammanställning har genererats med [ACE](#).

# DT2118 Mötesanteckningar (Niklas Vanhainen)

## Kurslitteratur

- En aspekt som togs upp var om kursboken ska vara obligatorisk. Anledningarna bakom detta var att ett flertal av kursdeltagarna inte använde sig av kursboken, samt att den är kostsam för institutionen.
- Om boken ändå används kommande år så är ett förslag att en mer detaljerad översikt över innehållet ges till kursdeltagarna. Detta kan vara bra då kursdeltagare har olika kunskapsbakgrunder och behöver hitta relevanta avsnitt.

## Artiklar

- Fler “review artiklar” är önskvärt.

## Kursinnehåll

- Viss repetition från kurs DT2112 förekommer.
- Bra föreläsningar. Fler föreläsningstillfällen är önskvärt. Exempelvis mer undervisning kring sökalgoritmer och språkmodeller.
- Kursdeltagarna upplevde att det var ett bra innehåll och nivå på hemuppgifterna.

## Tidsplan

- Tid till projekt / uppsats ansågs vara tillräcklig. Dock kan deadline för det första utkastet av projektet / uppsatsen tidigareläggas.
- Genom att dela upp hemuppgifterna i två inlämningstillfällen, varav den ena tidigareläggs, så kan detta bidra till att kursdeltagarna påbörjar uppgifterna tidigare.

## Svårighetsgrad på projekt

- Flera deltagare påbörjade projekt men fick ändra sina planer på grund av tidsbrist / felbedömning av svårighetsgrad. Även av denna anledning vore det bra om det första utkastet av projektet lämnades in tidigare för feedback.

## TIMIT

- Färdiga funktioner för TIMIT kan vara att föredra då det skulle underlätta uppstart av projekt där TIMIT används.

## Labb

- Konfiguration av HTK hade en tendens att ta en stor del av labbtiden. En lite tydligare lathund av kommandon för HTK samt shell script kan underlätta och flytta fokus från själva programmet till uppgiften.
- Att ha en gemensam labbtid för kursdeltagarna i en större sal kunde vara en potentiell förbättring.

## Betyg

- Att införa flergradig betygskala skulle kunna bidra till att fler studenter anmäler sig till kursen. Detta ur ett internationellt perspektiv då betyg kan behövas för vidare studier.
- Att delge delresultat till kursdeltagarna under kursens gång kan bidra till att kursdeltagarna får en ökad försåelse för hur dessa ligger till samt vilka områden som kan förbättras.

## Fler kursdeltagare

- Att synas på fler kurs- och programsidor kan troligen bidra till att fler studenter får upp ögonen för kursen.