Instructions for ANN fk projects

The projects should typically be carried out in groups of 2 students. The group should have consensus on the grade (A-B or C-E) to aim for. The final grade will be the same for all members of the group.

The project should be conducted in the spirit of a minor industrial R&D activity. The group is responsible for the project including background literature research, selection of or implementation of the necessary tools, and writing a report. Course teachers act as advisors and evaluators of the projects.

Project definition

In the course directory you will find a number of databases and brief task descriptions in separate sub-directories. These include data files and a file 'Definition.txt' describing briefly the data, the task, and also the criteria for a competition among the groups carrying out this task. Note that the projects/databases are quite different, for instance, with regard to the amount of data and the extent to which the data is ready for processing. Some databases require significant pre-processing before the actual processing can be started. Others are more or less well prepared to start this right away. A common component of all tasks is the competition that is defined in the Definition.txt (see also below).

Each group needs to compose a project from these databases and different types of ANN tools and possibly other suitable tools. There is considerable freedom in doing so and the project design should be given careful consideration by the group. As an exception, a group may request a different type of project than what is available in the course directory. That must then be discussed with and approved by a course teacher.

A grade A-B project should be quite ambitious, e.g. include substanital pre-processing of data (if necessary) and/or a comparison of, for instance, two different ANN tools and a third tool which in non-ANN. It could also include implementation of some component or tool. But note that it should also fit into the time allotted for the project (2 weeks/group participant)! Of course, a proper methodology should be used to find suitable parameters and achieve good generalization performance.

A grade C-E project may be more limited to the solution of one of the given tasks/databases with a standard NN tool, e.g. the NN toolbox of Matlab. It is, however, important that a proper methodology is used to find suitable parameters and that generalization is verified and a reasonable performance is reached.

A 1-2 A4 page specification of the project should be presented at the Project approval meeting (see time table below). It will be discussed, possibly modified, and then approved by a course teacher. At this time the group and the project will be properly registered and a contact person appointed.

Note that the specification should detail the planned distribution of tasks and resposibilities for each person in the project group.

Project report

Every project should be concluded with a project report. For a grade A-B project a 5 A4 pages high quality report is required. For the grade C-E projects a 3 A4 pages well written report is sufficient. The report should describe the project definition, databases and tools used, results reached, and a discussion of results and experiences reached during the course of the project.

Note that the project report should also detail the actual distribution of tasks and resposibilities for the persons in the project group.

Moreover, some of the projects will be selected for a 15 minutes oral presentation at the Final project seminar. This selection will be based on quality as well as to represent different types of projects and approaches.

Grading of project

The final grade for each category of project will be the same for all group members and will be based on to what degree the criteria for the respective category are fulfilled. If motivated, a grade A-B project can instead be given grades in the interval C-E. Grades are not dependent oh the competition results (below).

Project time table

Time	Room	Activity
Thu 2008-04-10, 13 – 14	RB35	Project kickoff
Thu 2008-04-17, 10 – 12	RB35	Proposed plan, acceptance
Wed 2008-05-07, 10 – 12	RB35	Project checkpoint
Wed 2008-05-14, at 12:59	N/A	Deadline for submission of competition entries
Thu 2008-05-15, 13 – 16	RB35	Final project seminar
Mon 2008-06-02	N/A	Deadline for project reports

At the Project kickoff the project phase of the course is initiated.

At the Proposed project specification time each group should present a 1-2 A4 project specification which is discussed with a course teacher. The database(s) utilized will be decided and a contact person for the project should be appointed by the group, the distribution of work is stated, etc. The final project specification is sent in (ala@nada.kth.se) and the group gets registered for the project.

At the Project checkpoint the group should be ready to present a brief first progress report. It will be checked that the project is on time and that the level of ambition is appropriate for the grade aimed for. There will be room for some advice from the course teachers.

ANN fk project – competition

Participation in the competition is not mandatory, but a stimulating and appreciated exercise! The competition is, however, not part of the course examination..

In summary, you submit an entry by sending an email containing a script that will execute your code. Given one or more input data files, this code should produce an output, e.g. a classification of the input test data items and a performance score. In the case of a time series, the set of data points that extends the time series (according to the description for the individual database) should be submitted. The Subject line should state the name of the database used and the name of the group's contact person.

Deadline for submission is Wednesday May 14, 12:59.

For time series tasks the extension of the time series is submitted. For classification tasks the following should be done:

1) Prepare a standalone version of our application that has been trained on one of the given databases. The infile will be 'test.indata' with the same format as the training data in the database (but with the label column removed if it appears in the training data file). The output should have the same format as in the training file, but nothing else than the predicted class label should be on a line in the output file.
2) Submit this code to ala@nada.kth.se before the deadline stated above. In case submissions are made for the same database, the submission closest before the deadline will be taken as the valid one.

For optimization tasks do the same as above, without the training part. The 'test.indata' file in this case will have city x,y positions, one per row. The output should be the order in which the cities are visited, the length of the tour, and a graphic display of the tour.

The entries submitted will be run and evaluated by the course teachers (judges) on the basis of:

1) That the code runs properly and produces an adequate output
2) The correctness of the output produced (for time series only 2).

The group that produced the application achieving the best result will be winner and in addition to the honour they will receive a small award. If there is just one entry for a database there will still be a winner, provided the application runs as expected and the result is reasonable.

Final project seminar and project report submission

At the Final project seminar the results and winners of the competitions will be presented. In addition, oral presentation of some projects and discussion and conclusions of the course projects will be organized.

Monday June 2 is the deadline for submitting the group's project report.