DIVISION OF STATISTICS AND MACHINE LEARNING
DEPT. OF COMPUTER AND INFORMATION SCIENCE

Associate professor
Oleg Sysoev
CAMPUS VALLA
SE-581 83 LINKÖPING
SWEDEN

Master's thesis in Statistics and Machine Learning in collaboration with industry

The two-year international Master's program **Statistics and Machine Learning** was started in 2008 at Linköping University to give students the necessary competence to tackle large-scale data problems, with special emphasis on prediction and decision making under uncertainty. It is a unique interdisciplinary program with students from statistics, computer science, engineering, and applied mathematics. The program includes core courses in statistics, machine learning, programming, database management and big data analytics.

A majority of the master theses in the program are written in collaboration with the industry, based on a specific problem that the firm wants to solve. This letter sets out the requirements for the thesis work as an aid for the firm. In particular, here are some important things to keep in mind while discussing a potential project:

- The thesis comprises 30 credit points, i.e. **100% studies for a semester** (the academic year in Sweden is divided in two semesters).
- The thesis is **organized as a course and runs between fixed dates** (Jan- June for the spring semester or Sept-Jan for the fall semester).
- The thesis **is individual work** and the student has to write a single-authored master thesis report.
- **The proposed project needs to be accepted** by the course examiner.
- The student will be assigned a main **supervisor** at the university. The role of the academic supervisor is to assist the student in **deciding on the methodology** for solving the given problem and to control the quality of the work.

- The firm needs to assign a contact person/external supervisor at the firm. **The firm is expected to contribute an applied problem formulation, supply the necessary data and to aid in evaluating the reasonableness of the student's results**.
- The thesis is in the subject of Statistics and machine learning applications **should therefore be based on statistically oriented methods.**
- The master **thesis is a scientific work**. This means for example that the proposed solution in the thesis needs to relate to existing work in the scientific literature, and the advantages and disadvantages of the proposed solution needs to be critically assessed.
- A master **thesis is a public document** the results from the thesis work will therefore be public and is to be presented in a public seminar. Any potential problems with making the findings public needs to be clearly discussed **before** the start of the thesis. In some special cases it may be possible to not reveal firm sensitive information, but this should be kept to an absolute minimum.

Oleg Sysoev, PhD
Program Director

LINKÖPING UNIVERSITY