

Ph.D. Program in
Machine Learning
Student Handbook





Ph.D. Program in Machine Learning

Student Handbook

Revised 8/20/14

Table of Contents

Introduction	3
Program Information	4
Program Requirements	4
Core Courses:	6
Electives:	7
Rules about previously taken courses:	7
The Research-Matching Process in ML	8
Role of the Advisor:	8
Data Analysis Project (DAP)	8
DAP Committee	9
DAP Prospectus	9
DAP Requirements:	9
Proficiencies in Programming, Teaching, Research and Writing Skills:	11
Machine Learning Journal Club	12
Directed Research	12
Ph.D. Proposal	13
All But Dissertation (ABD) Policy	14
Ph.D. Thesis	15
Student Evaluation	16
Financial Support	17
Travel Support	18
Leave of Absence Policy	18
Grievances	18
Seminars	19
The Emigration Course	19
General Information	20
Appendix A: University Resources	22

Introduction

The field of machine learning is concerned with the question of how computers can improve automatically through experience. Our Ph.D. program in Machine Learning is designed to give students a deep understanding of the computational and statistical principles that underlie learning processes, an exposure to real-world applications of machine learning, and an opportunity to design novel machine learning algorithms that advance the state of the art. As the only Machine Learning Department in existence, our goal is to produce graduates who go on to become leaders in this rapidly growing field. Our graduates have already gone on to take faculty positions in top-ranked Computer Science departments, Statistics departments, and Engineering departments at other universities, as well as positions in major industrial research laboratories.

The Ph.D. program is run by the Machine Learning Department which is part of Carnegie Mellon's School of Computer Science. This program builds on ML's world-class faculty, which includes a number of faculty with cross-appointments in diverse areas ranging from Statistics, Language Technologies, Philosophy, Psychology to the Tepper Business School.

Department Head of Machine Learning:

Tom Mitchell, Fredkin Professor of Artificial Intelligence and Learning.

Student Advising

The ML PhD program is supervised by two faculty co-directors. Graduate students can meet with these co-directors to discuss their curriculum or research. In addition each graduate student is matched with a faculty research advisor in the fall of the first year, who oversees the student's required Reading and Research course.

Co-Directors of the program:

Geoffrey Gordon, Associate Research Professor, Machine Learning Dept.
Email: (ggordon@cs.cmu.edu) Phone: x7399

Rob Kass, Professor, Statistics Dept.
Email: (kass@stat.cmu.edu) Phone: x8723

ML Department Administrative Support:

Diane Stidle, Graduate Programs Manager (diane@cs.cmu.edu) x1299

Russ O'Lare, Business Manager (rdo@cs.cmu.edu) x3384

Sharon Cavlovich, Assistant to the Dept. Head (sharonw@cs.cmu.edu) x5196

Alison Chiocchi, Employment Processes Manager (achioc@cs.cmu.edu) x9601

Mallory Deptola, Administrative Assistant (mdeptola@cs.cmu.edu) x5527

Sandra Winkler, Administrative Assistant (sandyw@cs.cmu.edu) x3569

Program Information

To complete the Ph.D. degree, we require that each student

- Participate in directed research
- In addition to research courses, pass 96 university units worth of graduate courses, with certain distribution requirements
- Complete a Data Analysis Project (DAP)
- Serve as a teaching assistant at least twice
- Demonstrate speaking, writing and programming skills
- Write and orally defend a thesis, a significant piece of original research in a specialized area of Machine Learning.

We are committed to the principle that students may achieve competence through a variety of methods, including courses, seminars, projects, and independent study. We consider each student's individual strengths, weaknesses, and interests in designing the best method for the student to fulfill these requirements. Our program is unique in that we encourage and expect students to engage in research from their first day in the Department.

Immigration Course (IC):

The Immigration course is intended primarily to introduce you to the Machine Learning Department and the School of Computer Science. Faculty give research presentations, with time for individual and group discussions. The research presentations will give you an idea of the kind of research that goes on here in Machine Learning and throughout SCS. The Immigration course is held either immediately before or at the beginning of the Fall term. After the IC, students are strongly encouraged to arrange individual meetings with potential faculty advisors.

Program Requirements

During the IC we will be discussing the courses you should register for during the first semester. As a full-time student you are required to register for 48 units each fall and spring semester. You will be conducting research with your advisor throughout the year and you must register for Reading & Research (10-920) each semester. This course is intended for you to work on research with your advisor, 50% of your time should be spent on research.

Prerequisites, Computer Science:

15-150 Principals of Functional Programming

An introduction to programming based on a "functional" model of computation. This course is an introduction to programming that is focused on the central concepts of function and type. One major theme is the interplay between inductive types, which are built up incrementally; recursive functions, which compute over inductive types by decomposition; and proof by structural induction, which is used to prove the correctness and time complexity of a recursive function. Another major theme is the role of types in structuring large programs into separate modules, and the integration of imperative programming through the introduction of data types whose values may be altered during computation. NOTE: students must achieve a C or better in order to use this course to satisfy the pre-requisite.

15-210 Parallel and Sequential Data Structures and Algorithms

Teaches students about how to design, analyze, and program algorithms and data structures. The course emphasizes parallel algorithms and analysis, and how sequential algorithms can be considered a special case. The course goes into more theoretical content on algorithm analysis than 15-122 and 15-150 while still including a significant programming component and covering a variety of practical applications such as problems in data analysis, graphics, text processing, and the computational sciences. NOTE: students must achieve a C or better in order to use this course to satisfy the pre-requisite.

Previously offered Computer Science courses 15-211 and 15-212 would also fulfill the prerequisite requirement.

Prerequisites, Statistics:

36-225: Introduction to Probability Theory

This course is the first half of a year-long course which provides an introduction to probability and mathematical statistics for students in economics, mathematics and statistics. The use of probability theory is illustrated with examples drawn from engineering, the sciences, and management. Topics include elementary probability theory, conditional probability and independence, random variables, distribution functions, joint and conditional distributions, law of large numbers, and the central limit theorem. A grade of C or better is required in order to advance to 36-226. Not open to students who have received credit for 36-625. 36-217 Probability Theory and Random Processes, will also be accepted as a prerequisite.

36-226: Introduction to Statistical Inference

This is mostly a theoretical course in statistics. First, we will give a formal introduction to point estimation and consider and evaluate different methods for finding statistical estimates. Then we will discuss interval estimation and hypothesis testing, which are necessary for most statistical analyses. In this first part of the course, the emphasis will be on definitions, theorems and mathematical calculations. Once we have covered the mathematical foundations of statistical inference, we will focus on the use of these concepts in concrete statistical situations. We will study statistical modeling and specific models such as ANOVA and regression. Emphasis will be placed on understanding the qualities of a good statistical analysis, specifying correct models, assessing model assumptions and interpreting results.

Previously offered Statistics courses 36-625 and 36-626 would also fulfill the prerequisite requirement.

Core Courses:

The three core courses listed below:

10-705: Intermediate Statistics

Some elementary concepts of statistics are reviewed, and the concepts of sufficiency, likelihood, and information are introduced. Several methods of estimation, such as maximum likelihood estimation and Bayes estimation, are studied, and some approaches to comparing different estimation procedures are discussed.

10-715: Advanced Introduction to Machine Learning

This course will give students a thorough grounding in the algorithms, mathematics, theories, and insights needed to do in-depth research and applications in machine learning. The topics of this course will in part parallel those covered in the general graduate machine learning course (10-701), but with a greater emphasis on depth in theory and algorithms. The course will also include additional advanced topics such as RKHS and representer theory, Bayesian nonparametrics, additional material on graphical models, manifolds and spectral graph theory, reinforcement learning and online learning, etc. Students entering the class are expected to have a pre-existing strong working knowledge of algorithms, linear algebra, probability, and statistics.

Note: Students who took 10-701 in Spring 2014 or earlier can use it as a core course, even if they weren't part of the MLD PhD program at the time they took 10-701.

10-702: Statistical Machine Learning

This course builds on the material presented in 10-701/10-715, introducing new learning methods and going more deeply into their statistical foundations and computational aspects. Applications and case studies from statistics and computing are used to illustrate each topic. Aspects of implementation and practice are also treated.

Plus any two of the following courses:

10-708 Probabilistic Graphical Models

This course will provide you with a strong foundation for both applying graphical models to complex problems and for addressing core research topics in graphical models. The class will cover three aspects: The core representation, including Bayesian and Markov networks, and dynamic Bayesian networks; probabilistic inference algorithms, both exact and approximate; and, learning methods for both the parameters and the structure of graphical models. Students entering the class should have a pre-existing working knowledge of probability, statistics, and algorithms, though the class has been designed to allow students with a strong numerate background to catch up and fully participate.

10-725 Convex Optimization

This course is designed to give a graduate-level student a thorough grounding in the formulation of optimization problems that exploit such structure, and in efficient solution methods for these problems. The main focus is on the formulation and solution of convex optimization problems.

15-826: Multimedia Databases and Data Mining

The course covers advanced algorithms for learning, analysis, data management and visualization of large datasets. Topics include indexing for text and DNA databases, searching medical and multimedia databases by content, fundamental signal processing methods, compression, fractals in databases, data mining, privacy and security issues, rule discovery and data visualization.

15-750 Graduate Algorithms

or

15-853 Algorithms in the Real World

This course covers how algorithms and theory are used in "real-world" applications. The course will cover both the theory behind the algorithms and case studies of how the theory is applied. It is organized by topics and the topics change from year to year.

Electives:

Electives may be chosen from Carnegie Mellon's large number of graduate courses, in consultation with the student's advisor & Co-Directors, to meet the interdisciplinary distribution requirements. One of these three electives is taken from the offerings in Statistics. The other two advanced electives, chosen in consultation with the student's advisor, form a research depth concentration in one of the allied disciplines with SCS, such as Biology, Statistics, or Tepper School of Business. For those candidates seeking an academic position after completing the ML Ph.D. degree, the thoughtful selection of these three elective courses is particularly important. As in each of the first two years, coursework is supplemented by 24 units/term of research.

Rules about previously taken courses:

If a student has taken some of the MLD core courses before joining the MLD PhD program, and has not counted these courses toward any other PhD-level degree, the student may count these courses toward the MLD PhD. In this situation the student will need to take fewer than 5 new core courses to graduate. A student must always take at least three elective courses while registered in the MLD PhD program, irrespective of any courses taken before joining the PhD program. Students who took 10-701 in Spring 2014 or earlier can use it as a core course, even if they weren't part of the MLD PhD program at the time they took 10-701.

If any courses are counted toward both the MLD PhD and a lower-level degree (such as an MS) from another department outside of MLD, the student will not receive an MS degree from MLD. If the student wishes to receive an MLD MS in addition to the outside degree, s/he must take additional courses so as not to count any courses toward both the MLD MS degree and the outside degree.

The Research-Matching Process in ML

Carnegie Mellon is a research institution. We are strongly committed to scientific excellence, both in research and education. In particular, we believe that a close personal interaction among students, faculty, and staff is of the utmost importance for educating the next generation of leaders in academia and industry. ML students are therefore matched to a faculty advisor in the very beginning of the program who will guide their research and advise them in academic matters.

Approximately two weeks after the end of the IC, both ML students and Machine Learning (and affiliated) faculty submit a form, indicating preferences for advisorship relations. Based on these forms, the Co-Directors of the ML program will then match students with faculty advisors. Each student either will be assigned to one advisor, or will be co-advised by two faculty advisors. A student's advisor may change if the research direction changes and there is no longer a match.

Role of the Advisor:

The faculty advisor is a student's primary contact, both in research and in academic matters. Typically, a student has strong interests in the research area of the faculty advisor, and she/he will closely collaborate with the faculty member. The advisor is typically the primary person directing the student research, and is also expected to provide financial support (stipend and tuition) for the student.

Data Analysis Project (DAP)

Students are required to demonstrate their grasp of fundamental data analysis and machine learning concepts and techniques in the context of a focused project. The project should focus on a substantive problem involving the analysis of one or more data sets and the application of state-of-the-art machine learning and data mining methods, or on suitable simulations where this is deemed appropriate. Or, the project may focus on machine learning methodology and demonstrate its applicability to substantial examples from the relevant literature. The project may involve the development of new methodology or extensions to existing methodology, but this is not a requirement.

Machine learning and data mining methods are exemplified by, but not limited to, those covered in the core courses 10-701, 10-702, and 15-826. In particular, the analysis methods should be adequately justified in terms of the theory taught in these courses.

The project is not intended for purely theoretical or methodological investigations, but these may form the heart of a project in appropriate cases. (In such cases, the project should also contain a component of applying the new theoretical or methodological tools to data. This component does not have to contain novel results; instead, its goal is to characterize how well or poorly the tools perform for the given data.) Students are encouraged to seek out a project (co)advisor who can provide access to data or substantive applications, or can use data sets to which they already have access through one of the core courses, through the literature and archives, or through their PhD advisor. Other resources for this purpose include the Immigration Course, faculty home pages, and the ML Research Projects webpage.

The Data Analysis Project is to be carried out under the supervision of a Machine Learning Department faculty member, and possibly under joint supervision of a subject matter expert. It is to be concluded by a written report. The ideal report would demonstrate an ability to approach machine learning problems in a way that cuts across existing disciplinary boundaries. It should demonstrate a capacity to write about technical topics in machine learning in a cogent and clear manner for a professional and scientific audience.

Research for the Data Analysis Project is typically done as part of the Reading & Research course, 10-920. The student must register for the ML Journal Club, 10-915, for the semester they intend to present their Data Analysis project.

DAP Committee

Student must form an official "DAP committee" of three faculty to evaluate the document. The committee will consist of the advisor, the Journal club instructor(s), and one other faculty member selected by the student. The third member is often someone with an interest in the analysis of the data set, and does not have to be an expert in ML or part of the student's thesis committee. The student should form the committee as early as possible during the DAP research process, and inform Diane of who the members are. 2 of 3 DAP Committee members, one of whom is the DAP advisor, must be in attendance for the DAP presentation.

DAP Prospectus

Student must write a 1-2 page prospectus, including the DAP's title, general topic, proposed data source, and a brief summary of proposed analysis methods, and circulate it to the committee. The student should do this as early as possible, preferably when the student forms the committee.

The intent is that the Data Analysis Project will be less formal in structure and more flexible in focus than a typical Masters thesis + defense requirement might allow. The Project is a requirement for those in other departments receiving a MS degree in Machine Learning as well as for PhD students in Machine Learning. The requirement will typically be completed during a student's 2nd year in the program.

DAP Requirements:

- 1) A presentation of the work during the Machine Learning Journal Club course. The presentation stands in lieu of a defense of the Data Analysis Project, and helps to disseminate the work to the rest of the Machine Learning community. There will be a limited set of dates available for such presentations---generally, at most one per week---so students should be sure to sign up early in the Machine Learning Journal Club. The presentation should be suitable for a general machine learning audience, i.e., it should provide sufficient background for a non-domain-expert to understand the results, and should adequately summarize the relationship of the project to previous work. 2 of 3 DAP Committee members, one of whom is the DAP advisor, must be in attendance.

- 2) A stand-alone, single or “lead author” written paper that is approved by the faculty member(s) advising the Project. The paper should be of high quality, both in terms of exposition of technical details and overall English and organization. It should be suitable for submission to a journal or refereed conference. But, unlike some conference papers, it should be completely self-contained, including all descriptions necessary for a general machine learning audience to follow the theoretical development and reproduce the experimental results. This requirement may (but does not have to) result in the project paper being substantially longer than a conference proceedings paper on which it is based. Although it does not have to be published, publishing the paper may be desirable and helpful to the student. Project papers will become part of the MLD archives, and will serve as examples to future students.
- 3) The student must provide a near-final draft of the DAP document (approximately 15 pages) at least one month before the oral presentation to the DAP Committee. Both student and committee must certify that this draft is substantially complete. Within two weeks of submission, the instructor(s) will either approve the project for presentation (at which point the presentation can be advertised to the members of the department), or notify the student that changes will be required before presentation. This approval is for the general topic and content, and not for the final contents of the document. The final version of the paper, incorporating any feedback received at the oral presentation, should be submitted for review no later than one month after the oral presentation.

Proficiencies in Programming, Teaching, Research and Writing Skills:

- The programming skill requirement is normally demonstrated during the student's first two years of research, carried out under the supervision of the student's research advisor.
- Each Ph.D. candidate must participate in two terms of instruction, either through TA duties or serving as the instructor for a class.
- Research and writing skills are normally achieved through the Data Analysis Project requirement.

Programming Requirement:

The programming requirement is that a faculty member vouch for the student's ability to build a reasonably sized system with reasonably clean code. This ability includes designing the system well, explaining its operation to others (to the level that they could use it, e.g., via a code release), and if applicable collaborating with others to develop it. The faculty member should have personal knowledge of this ability before certifying it, e.g., based on a code review, or on detailed discussions of the design of the system coupled with observations of a successful code release.

Speaking Requirement:

To satisfy the oral communication skill requirement, each student should give a public talk at Carnegie Mellon. The talk should be accessible to a general SCS audience. The talk is presented in the Machine Learning Journal Club, and members of the speaking committee attend and evaluate the presentation, as well as provide oral and written feedback to the student.

All members of the speaking committee and the student's advisor should be in attendance. Immediately after the talk, the speaking committee members and the student's advisor confer among themselves (with the student absent) about the presentation. The committee members also fill out a Speaking Skills Review Form. The committee then provides oral feedback to the student, and give the review forms to the student. After studying their contents, the student submits the review forms to the Graduate Program Administrator to receive credit and to have placed in the Student's file.

As with writing, speaking well takes practice. Satisfying this requirement might take a few tries on the student's part, and no stigma is attached to those who have to try more than once. They are, however, encouraged to satisfy this requirement as soon as possible, and should consult with their advisor to choose an appropriate schedule, preferably during the first year.

Machine Learning Journal Club

10-915 the ML Journal Club:

Course website: <http://www.cs.cmu.edu/~journalclub/>

The ML Journal Club is a once-a-week course designed to give students experience in making presentations. All students are required to take the Journal Club course twice, typically over the course of their first three years. In addition to practicing presentation skills, you will use the Journal Club to satisfy two requirements: the speaking skills requirement and the Data Analysis Project oral requirement. Typically, students satisfy speaking skills during their first year and the DAP oral during their second or third year, but this schedule might vary depending on how much presentation experience someone has before joining the MLD. New students should register in the fall or spring semester of their first year to attempt to satisfy the speaking skills requirement.

Sign up in advance to schedule your talk

We will open up the sign-up sheet for talk slots in advance of the course start date: you must sign up for a slot in order to register for the course. Those students who have already taken 10-915 twice and still need to finish a talk requirement must sign up in advance for a talk but are not required to register for a third time.

Advisor Attendance

Advisors are to attend both the student's speaking skills talk and the DAP oral. Student must check with their Advisor to make sure they will attend.

Student Attendance

Students are required to attend all lectures in order to pass, unless they get permission from the instructor(s) to skip (a small number of) lectures due to travel, etc.

Directed Research

During a student's first two years, he or she should be doing directed research at least half time; once all coursework is completed, full time (except when teaching). Different students, and different advisors, have different ideas of what directed research means and how progress can be demonstrated. It is the responsibility of both the student and his or her advisor to formulate for each semester a set of reasonable goals, plans, and criteria for success in conducting directed research. Advisors are individually responsible for adequately supervising this portion of the Ph.D. program.

Ph.D. Proposal

ML PhD students are expected to present their thesis proposal during their third or fourth year. Typically, the proposal is completed by the beginning of the fourth year. Fulfilling the requirement involves writing and orally presenting a proposal, and obtaining advice and approval from the thesis committee. Students should meet with the committee members at least once to discuss the proposed work before the proposal. The thesis committee should be composed of four or more members (including the student's research advisor), at least one of whom is an external member and at least one of whom is a Machine Learning faculty member. The external member is typically from outside the university but could be from another department at Carnegie Mellon if appropriate. The Department has the right in unusual cases to alter the composition of the committee to assure appropriate quality and breadth.

Students should allocate at least 2.5 hrs. for the proposal presentation and examination. In addition to the student, at least two committee members, one of whom is the Chair, must be physically present at CMU for the proposal; other committee members may attend by teleconference.

Generally, a thesis proposal will be approximately 15 pages plus references, and will include (a) a clear statement of the research problem and proposed research, (b) a discussion of related research and how the proposed work fits into the field, (c) a description of the technical approach, (d) preliminary research results that demonstrate the proposed research is plausible and worthwhile, (e) a discussion of research issues to be pursued, and (f) a tentative schedule for completing the work. Of course in a proposal it is impossible to predict precisely which research issues will be solved in the future. Nevertheless the proposal should include a list of specific research directions and questions that are likely to be addressed, and for each of these an assessment indicating what could be a baseline approach, and a discussion of ideas for pursuing the issue, along with an assessment of what will be easy versus difficult. The student needs to show that the proposed research will be original and interesting, and that it is likely to succeed. During the later thesis defense, the student will *not* be required to show that he or she has done everything that was proposed. In this sense, the proposal is an opportunity to present the student's best current ideas about the thesis research, and obtain some useful early feedback from experts in the research area. The proposal need not have answers to every question it raises, but it should bring up a good list of questions that will drive the research.

If you would like to see examples of some previous thesis proposals, please contact your Graduate Program Administrator.

All But Dissertation (ABD) Policy

After the presentation of an acceptable thesis proposal, and satisfying all other requirements except for the dissertation and its oral defense, students are regarded as "all but dissertation."

Time to Degree:

Once students achieve ABD status, students who began in the PhD program prior to June 1, 2011 must complete all requirements for the PhD within a maximum of seven full academic years, unless terminated earlier by conferral of the degree or by academic or administrative action.

Students who began in the PhD program after June 1, 2011 must complete all requirements for the PhD within a maximum of ten years from original matriculation as a doctoral student, unless terminated earlier by conferral of the degree or by academic or administrative action.

Once this time-to-degree limit has lapsed, the person may resume work towards a doctoral degree only if newly admitted to a currently offered doctoral degree program under criteria determined by that program.

An ABD candidate may choose to continue as a regular student *In Residence*, or to be *In Absentia* (ABS).

Please see the University policy: <http://www.cmu.edu/policies/DSS.html>

Ph.D. Thesis

Normally, the thesis dissertation is completed during the student's fifth year. The thesis must describe a significant piece of original research work. It is evidence of proficiency, high attainment, and ability to do research in a specialized area of Machine Learning.

The final defense is a public presentation, in accord with the College and University requirements for the Ph.D. It is the candidate's responsibility to ensure that the College and University's guidelines are followed for publicity of the defense, and the availability of the thesis at least one week prior to the defense.

Work with the Graduate Program Administrator to determine timing so as to avoid department and class conflicts. Contact your thesis committee to get their availability. You should send a draft of the thesis to your committee about one month before you plan to defend. Your committee should get back to you with their approval to defend before the announcement goes out, two weeks before your defense date.

Students should allocate at least 3 hours for the thesis defense and examination. In addition to the student, at least three committee members, one of whom is the Chair, must be physically present at CMU for the defense. All committee members must either be in attendance or attend by teleconference.

The presentation by the candidate is normally about 45 minutes. The thesis committee chair (advisor) determines who may ask questions and in what order and brings the discussion to a close at the appropriate time. The question-and-answer period is followed by a closed-door session attended by only the members of the thesis committee and any interested faculty members. If the student passes the oral presentation the options of the committee are:

To approve without corrections

To approve subject to minor changes, to be approved later by the thesis chair only

To require a resubmission after major changes and re-approval of the entire committee

Not to approve the thesis

All members of the committee are required to sign a Final Oral Examination card to indicate that the student has passed the thesis oral examination.

In addition, the thesis committee chair, the Department Head, and the Dean sign a final certification sheet after final approval of the thesis by the thesis committee and student has submitted the final version to the Graduate Programs Administrator.

Student Evaluation

The faculty meet at the end of each academic semester to make a formal evaluation of each student in the program. For historical reasons this meeting is called "Black Friday." The co-directors and faculty research advisors communicate in written and oral form the assessment from these Black Friday meetings to the graduate students.

Evaluation and feedback on a student's progress are important both to the student and to the faculty. Students need information on their overall progress to make long range plans.

At each semi-annual "Black Friday" meeting, the faculty review the student's previous semester's research progress and the student's next semester's research plans to ensure that the student is making satisfactory progress. The evaluation of a student's progress in directed research often depends on the student having produced some tangible result; examples include the implementation of pieces of a software system, a written report on research explorations, an annotated bibliography in a major area, or, as part of preparation for doing research, a passing grade in a graduate course (beyond the required 96 required units).

The purpose of having all the faculty meet together to discuss all the students is to ensure uniformity and consistency in the evaluation by all of the different advisors. The faculty measure each student's progress against the goal of completing the program in a reasonable period of time. In their evaluation the faculty consider courses taken, directed research, teaching if applicable, skill, development, papers written and lectures.

The faculty's primary source of information about the student is the student's advisor. The advisor is responsible for assembling the above information and presenting it at the faculty meeting. The student should make sure the advisor is informed about participation in activities and research progress made during the semester. Each student is asked to submit a summary of this information to the advisor at the end of each semester.

Based on the above information, the faculty decide whether a student is making satisfactory progress in the program. If so, the faculty usually suggest goals for the student to achieve over the next semester. If not, the faculty make more rigid demands of the student.

Ultimately, permission to continue in the program is contingent on whether or not the student continues to make satisfactory progress in their home department and toward the ML degree. If a student is not making satisfactory progress, the faculty may choose to drop the student from the program.

Terms of progress in Black Friday letters from faculty:

SP = In the semiannual evaluation of all our students the faculty reviewed your progress toward the Ph.D. We are happy to report that you are in good standing in the Machine Learning PhD program.

USP = We have determined that your current level of progress is unsatisfactory:

N-2 = We have determined that there are significant problems with your current level of progress. Accordingly, this is an N-2 letter: you are in danger of receiving an N-1 letter next Black Friday unless you improve your rate of progress toward a Ph. D. In particular:

N-1 = This is an N-1 letter. You may not be allowed to continue in the PhD program past the next Black Friday meeting unless you satisfy the following conditions:

Financial Support

The Machine Learning Department is committed to providing full tuition and stipend support for the academic year, for each full-time ML Ph.D. student, for a period of 5 years. Research opportunities are constrained by funding availability. The funding commitments assume that the student is making satisfactory progress in the program, as reported to the student at the end of each academic term. Students are strongly encouraged to compete for outside fellowships and other sources of financial support. The department will supplement these outside awards in order to fulfill its obligations for tuition and stipend support.

Travel Support

The department encourages students to travel to conferences and workshops to enhance their professional and career development.

Policy: If a student wants to attend a conference or workshop, the student's advisor or research sponsor should support the trip through either a research contract or a discretionary account. Student travel is unlimited as long as there is money available from research contracts and/or discretionary funds of a sponsoring faculty member.

If no such funding is available to the student, then limited departmental funds may be available upon request from the Machine Learning Department. Since departmental funds are limited, the maximum to be reimbursed will be \$200 plus the registration fee, if only attending the conference or workshop; \$600 plus registration fee, if presenting a paper. Department funding is only available to the student for one trip per year and will not be transferred to the following year. This funding is only available if the Advisor agrees with the student's decision to attend the conference but does not have the funds.

Process: To obtain travel support, the student and his or her faculty advisor/research sponsor must first agree that the student should take the trip. Then in advance of the trip the student should print the Student Travel Authorization Form, and then get the advisor's signature before forwarding the form to the PhD Program Administrator. The faculty member must either (i) indicate the amount of support the student may receive and its source (be sure the charge number is filled in!), or (ii) state on the Comments line that no funds are available from any research or discretionary account. Request a Travel Form from the Graduate Programs Administrator.

Leave of Absence Policy

Students who wish to leave the program temporarily may request a leave of absence by submitting a request to the PhD Administrator. Leaves are initially granted for a period of no more than one year, but an extension of up to one additional year may be granted under exceptional circumstances. When an extension is granted, the conditions for return must be negotiated with the advisor and the Ph.D. Program Co-Directors, prior to returning to the program.

Students on leave of absence should contact the PhD Program Administrator two months prior to the end of the leave to indicate their plans for the next year.

Grievances

In case of grievances, the Machine Learning Department follows University grievance procedures; please refer to those procedures for more information.

<http://www.cmu.edu/graduate/policies/Summary%20of%20Graduate%20Student%20Appeal%20and%20Grievance%20Procedures.html>

Seminars

The Machine Learning Department sponsors seminars by researchers from within and outside Carnegie Mellon, which are attended by faculty, staff and graduate students. Students are encouraged to meet and interact with visiting scholars. This is extremely important, both to get a sense of the academic projects that are pursued outside of Carnegie Mellon and to get to know the leaders of such projects. That applies not only to seminars directly relevant to a student's research interests: the seminars provide an opportunity to widen one's perspective on the field.

The Emigration Course

The Emigration Course grooms finishing students for their career afterward. It is structured as a series of talks offered throughout the year and focuses on five topics: Jobs, The Real World, Money, Ethics, and Communication. These talks cover nuts-and-bolts issues like how to job interview, how to apply for grant money, and how to write a technical paper. They also expose students to traditional and non-traditional career paths in academia, industry, and government. Participation is open to the entire SCS community and is completely voluntary. More senior students, especially those planning to finish in any given year, are encouraged to attend sessions offered that year; however, even junior students can benefit from attending, to prepare for a smooth transition from life as a student to life in the real world.

General Information

Computers

Every incoming ML PhD student will have a computer on his or her desk. All ML PhD students will be given a CS computer account. The School of Computer Science has a Help Center located at 4203 Gates Building. If you have any computer or account problems contact help@cs.cmu.edu or call 8x4231 from a campus phone.

Photocopier/Printers

In order to use the photocopiers you must have an access code. In order to use the Black & white and color copiers, you must have a code #9136 for scanning and copies but not for printing. The copier codes are to be used by Machine Learning students only and are not to be given out to anyone not currently in the program.

To find out other names/locations of printers, please see the SCS Facilities website: <http://www.cs.cmu.edu/~help/printing/>

CS Main Office

We do not have our own Main Office or mail facilities. CS is allowing us to use the following services.

Functions of the CS Main Office

- Send mail
- Pick up mail/packages
- Send overnight packages

Your mail will be in the Gates Building, 6th floor.

Please do not take any supplies from the CS Main office, you are to get them from your Graduate Programs Administrator.

To have packages delivered to you please use the following address:

Your Name
Machine Learning Dept.
School of Computer Science
6105 Gates Building
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

US Post Office is located in the basement of University Center.

Appendix A

Highlighted University Resources for Graduate Students and The WORD, Student Handbook

Key Offices for Graduate Student Support

Office of the Assistant Vice Provost for Graduate Education

www.cmu.edu/graduate; grad-ed@cmu.edu

The Office of the Assistant Vice Provost for Graduate Education, AVPGE, directed by Suzie Laurich-McIntyre, Assistant Vice Provost for Graduate Education, provides central support for graduate students in a number of roles. These include: being an ombudsperson and resource person for graduate students as an informal advisor; resolving formal and informal graduate student appeals; informing and assisting in forming policy and procedures relevant to graduate students; and working with departments on issues related to graduate students and implementation of programs in support of graduate student development.

The Office of the AVPGE often partners with the division of Student Affairs to assist graduate students with their Carnegie Mellon experience. Senior members of the student affairs staff are assigned to each college (college liaisons) and are often consulted by the Assistant Vice Provost for Graduate Education and departments on an individual basis to respond to graduate student needs.

The Office of the Assistant Vice Provost for Graduate Education (AVPGE) offers a robust schedule of professional development opportunities. Some are geared towards a specific population (master's students, PhD students at the beginning of their program, graduate students seeking tenure track positions, etc.) and others are open to all graduate students (time management, balancing, staying healthy). A full schedule of programs can be found at: <http://www.cmu.edu/graduate/>.

The Office of the AVPGE also coordinates several funding programs, and academically focused seminars and workshops that advise, empower and help retain all graduate students, particularly graduate students of color and women in the science and technical fields. The fundamental goals of our programs have been constant: first, to support, advise and guide individual graduate students as they work to complete their degrees; second, to contribute to the greatest degree possible to the diversification of the academy. Visit the Graduate Education website for information about:

- Conference Funding Grants
- Graduate Small Project Help (GuSH) Research Funding
- Graduate Student Professional Development: seminars, workshops and resources
- Graduate Women Gatherings (GWG)
- Inter-university Graduate Student of Color Series (SOC)

Office of the Dean of Student Affairs

www.cmu.edu/student-affairs/index.html

The Office of the Dean provides central leadership of the metacurricular experience at Carnegie Mellon. The offices that fall under the division of Student Affairs led by Dean of Student Affairs Gina Casalegno, include:

- *Career and Professional Development Center*
- *Counseling & Psychological Services (CAPS)*
- *Housing & Dining Services*
- *Orientation & First Year Programs (note: for undergraduate students)*
- *Office of International Education (OIE)*
- *Student Activities*
- *Student Life.*

Holly Hippensteel, Assistant Dean of Student Affairs, serves as the point person in the division for graduate student resources and concerns. Graduate students will find the enrollment information for Domestic Partner Registration in the Office of the Dean of Student Affairs and on the website. The Office of the Dean of Student Affairs also manages the Emergency Student Loan (ESLs) process. The Emergency Student Loan service is made available through the generous gifts of alumni and friends of the university. The Emergency Student Loan is an interest-free, emergency-based loan repayable within 30 days. Loans are available to enrolled students for academic supplies, medication, food or other expenses not able to be met due to unforeseeable circumstances. The Office of the Dean of Student Affairs also provides consultation, support, resources and follow-up on questions and issues of Academic Integrity: www.cmu.edu/academic-integrity.

Assistance for Individuals with Disabilities

Students with disabilities are encouraged to self-identify with Equal Opportunity Services by contacting Larry Powell, 412-268-2013, lpowell@andrew.cmu.edu to access the services available at the university and initiate a request for accommodations.

Eberly Center for Teaching Excellence & Educational Innovation

www.cmu.edu/teaching

Support for graduate students who are or will be teaching is provided in many departments and centrally by the Eberly Center for Teaching Excellence & Educational Innovation. The Eberly Center offers activities for current and prospective teaching assistants as well as any graduate students who wish to prepare for the teaching component of an academic career. The Center also assists departments in creating and conducting programs to meet the specific needs of students in their programs. Specific information about Eberly Center support for graduate students can be found at: www.cmu.edu/teaching/graduatestudentsupport/index.html.

Graduate Student Assembly

www.cmu.edu/stugov/gsa/index.html

The Carnegie Mellon Student Government consists of an Executive Branch and a Legislative Branch. This is the core of traditional student government, as governed by the Student Body Constitution. The Executive Branch serves the entire student body, graduate and undergraduate, and consists of one president and four vice-presidents. The Legislative Branch for graduate students, The Graduate Student Assembly (GSA)

passes legislation, allocates student activities funding, and otherwise acts on behalf of all graduate student interests. GSA also plans various social opportunities for graduate students and maintains a website of graduate student resources on and off-campus, www.cmu.edu/stugov/gsa/resources/index.html. Each department has representation on GSA and the department rep(s) is the main avenue of graduate student representation of and information back to the graduate students in the department.

Intercultural Communication Center (ICC)

www.cmu.edu/icc/

The Intercultural Communication Center (ICC) is a support service offering both credit and non-credit classes, workshops, and individual appointments designed to equip nonnative English speakers (international students as well as international students who attended high school in the U.S.) with the skills needed to succeed in academic programs at Carnegie Mellon. In addition to developing academic literacy skills such as speaking, reading and writing, students can learn more about the culture and customs of the U.S. classroom. The ICC also helps international teaching assistants (ITAs) who are non-native English speakers develop fluency and cultural understanding to teach successfully at Carnegie Mellon and provides ITA testing.

Office of International Education (OIE)

www.studentaffairs.cmu.edu/oie/

Carnegie Mellon hosts international graduate and undergraduate students who come from more than 90 countries. Office of International Education (OIE) is the liaison to the University for all non-immigrant students and scholars. OIE provides many services including: advising on personal, immigration, academic, social and acculturation issues; presenting programs of interest such as international career workshops, tax workshops, and cross-cultural and immigration workshops; supporting international and cultural student groups such as the International Student Union and the International Spouses and Partners Organization; maintaining a resource library that includes information on cultural adjustment, international education and statistics on international students in the United States; posting pertinent information to students through email and the OIE website, and conducting orientation programs.

Key Offices for Academic & Research Support

Computing and Information Resources

www.cmu.edu/computing

Computing Services provides a comprehensive computing environment at Carnegie Mellon. Graduate students should seek Computing Services for information and assistance with your Andrew account, network access, computing off-campus, campus licensed software, email, calendar, mobile devices, computer security, cluster services and printing.

The Carnegie Mellon Computing Policy establishes guidelines and expectations for the use of computing, telephone and information resources on campus. The policy is supported by a number of guidelines graduate students should know. The policy and guidelines are available at: www.cmu.edu/computing/guideline/index.html.

Research at CMU

www.cmu.edu/research/index.shtml

The primary purpose of research at the university is the advancement of knowledge in all fields in which the university is active. Research is regarded as one of the university's major contributions to society and as an essential element in education, particularly at the graduate level and in faculty development. Research activities are governed by several university policies. Guidance and more general information is found by visiting the Research at Carnegie Mellon website.

Office of Research Integrity & Compliance

www.cmu.edu/research-compliance/index.html

The Office of Research Integrity & Compliance (ORIC) is designed to support research at Carnegie Mellon University. The staff work with researchers to ensure research is conducted with integrity and in accordance with federal and Pennsylvania regulation. ORIC assists researchers with human subject research, conflicts of interest, responsible conduct of research, export controls, intellectual property rights and regulations, and institutional animal care & use. ORIC also consults on, advises about and handles allegations of research misconduct.

Key Offices for Health, Wellness & Safety

Counseling & Psychological Services

www.studentaffairs.cmu.edu/counseling

Counseling & Psychological Services (CAPS) affords the opportunity for students to talk privately about issues that are significant for them in a safe, confidential setting. Students sometimes feel confused about why they are feeling upset and perhaps confused about how to deal with it. An initial consultation with a CAPS therapist will clarify options and provide a recommendation to the appropriate mental health resource at Carnegie Mellon or the larger Pittsburgh community. CAPS services are provided at no cost. Appointments can be made in person or by telephone, 412-268-2922.

Health Services

www.cmu.edu/HealthServices/

University Health Services (UHS) is staffed by physicians, advanced practice clinicians and registered nurses who provide general medical care, allergy injections, first aid, gynecological care and contraception as well as on-site pharmaceuticals. There is a small visit fee to see the physicians and advanced practice clinicians; nurse visits are free of charge. Fees for prescription medications, laboratory tests, diagnostic procedures and referral to the emergency room or specialists are the student's responsibility. UHS also has a registered dietician and health promotion specialists on staff to assist students in addressing nutrition, drug and alcohol and other healthy lifestyle issues. In addition to providing direct health care, UHS administers the Student Health Insurance Program. The Student Health Insurance plan offers a high level of coverage in a wide network of health care providers and hospitals. It also covers most of the fees for care at Student Health Services. Graduate students should contact UHS to discuss options for health insurance for spouses, domestic partners and dependents. Appointments can be made by visiting UHS's website or by telephone, 412-268-2157.

University Police

<http://www.cmu.edu/police/>

412-268-2323 (emergency only), 412-268-6232 (non-emergency)

The University Police Department is located at 300 South Craig Street, Room 199 (entrance is on Filmore Street). The department's services include police patrols and call response, criminal investigations, shuttle and escort services (additional information included in the Parking and Transportation section of The WORD, see below), fixed officer and foot officer patrols, event security, and crime prevention and education programming. Visit the department's website for additional information about the staff, escort and shuttle, emergency phone locations, crime prevention, lost and found, finger print services, and annual statistic reports.

Carnegie Mellon University publishes an annual campus security and fire safety report describing the university's security, alcohol and drug, sexual assault, and fire safety policies and containing statistics about the number and type of crimes committed on the campus and the number and cause of fires in campus residence facilities during the preceding three years. Graduate students can obtain a copy by contacting the University Police Department at 412-268-6232. The annual security and fire safety report is also available online at www.cmu.edu/police/annualreports.

The WORD

<http://www.cmu.edu/student-affairs/theword/>

The WORD is Carnegie Mellon University's student on-line handbook and is considered a supplement to the department (and sometimes college) handbook. The WORD contains campus resources and opportunities, academic policy information and resources, community standards information and resources. It is designed to provide all students with the tools, guidance, and insights to help you achieve their full potential as a member of the Carnegie Mellon community. Information about the following is included in The WORD (not an exhaustive list) and graduate students are encouraged to bookmark this site and refer to it often:

Carnegie Mellon Vision, Mission

Carnegie Code

Academic Standards, Policies and Procedures

 Educational Goals

 Academic and Individual Freedom

 Statement on Academic Integrity

 Standards for Academic & Creative Life

 Assistance for Individuals with Disabilities

 Master's Student Statute of Limitations

 Conduct of Classes

 Copyright Policy

 Cross-college & University Registration

 Doctoral Student Status Policy

 Evaluation & Certification of English Fluency for Instructors

 Final Exams for Graduate Courses

 Grading Policies

 Intellectual Property Policy

 Privacy Rights of Students

 Research

 Human Subjects in Research

 Office of Research Integrity & Compliance

 Office of Sponsored Programs

 Policy for Handling Alleged Misconduct of Research

Policy on Restricted Research
Student's Rights
Tax Status of Graduate Student Awards

Campus Resources & Opportunities

- Alumni Relations
- Assistance for Individuals with Disabilities
- Athletics, Physical Fitness & Recreation
- Carnegie Mellon ID Cards and Services
- Cohon University Center
- Copying, Printing & Mailing
- Division of Student Affairs
- Domestic Partner Registration
- Emergency Student Loan Program
- Gender Programs & Resources
- Health Services
- Dining Services
- The HUB Student Services Center
- ID Card Services
- Leonard Gelfand Center
- LGBTQ Resources
- Multicultural and Diversity Initiatives
- Opportunities for Involvement
- Parking and Transportation Services
- SafeWalk
- Survivor Support Network
- Shuttle and Escort Services
- Spiritual Development
- University Police
- Student Activities
- University Stores

Community Standards, Policies and Procedures

- Alcohol and Drugs Policy
- AIDS Policy
- Bicycle/Wheeled Transportation Policy
- Damage to Carnegie Mellon Property
- Deadly Weapons
- Discriminatory Harassment
- Disorderly Conduct
- Equal Opportunity/Affirmative Action Policy
- Freedom of Expression Policy
- Health Insurance Policy
- Immunization Policy
- Missing Student Protocol
- Non-Discrimination Policy
- On-Campus Emergencies
- Pets
- Political Activities
- Recycling Policy
- Riotous and Disorderly Behavior
- Safety Hazards

Scheduling and Use of University Facilities
Sexual Harassment and Sexual Assault Policy
Smoking Policy
Student Accounts Receivable and Collection Policy and Procedures
Student Activities Fee
Student Enterprises
Workplace Threats and Violence Policy
Statement of Assurance

University policies can also be found in full text at: <http://www.cmu.edu/policies/>

Last updated: June 24, 2014



MACHINE LEARNING
DEPARTMENT

School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
www.ml.cmu.edu