THE UNIVERSITY of EDINBURGH
informatics

Course Proposal Form

Please see Page 2 for instructions on which parts of this form to complete, whom to consult with to avoid unnecessary effort, and where to send the completed form.

Proposer(s): Rik Sarkar **Date: 12 Nov. 2019**

Cover page: Basic permanent course information

Unless otherwise noted, items in this section are entered into EUCLID and **cannot** be changed without creating an entirely new course.

Course Name	Machine Learning Theory
Course Acronym *(used by the School only, e.g., for the Sortable Course List)*	MLT
Course Level If the course is **only** available to MSc students, then it must be classed as Postgraduate. All other courses, regardless of level, are Undergraduate.	_x_ Undergraduate __ Postgraduate
Normal Year Taken	__UG1 __UG2 __UG3 _x_ UG4 __UG5 __MSc
Also available in years *[This can be changed later if need be.]*	__UG1 __UG2 __UG3 __UG4 _x_ UG5 _x_ MSc
SCQF Credit Level Level 8 should normally be used for pre-honours courses. Level 10 should normally be used for optional UG3 courses (so UG4 students may also take them) and for courses aimed mainly at UG4 students. Level 11 should be used for courses aimed mainly at MSc students, whether or not UG4 students can also take them.	__7 __8 __9 __10 _x_ 11
SCQF Credit Points	_x_ 10 __20 __40 __60 __80 __Other:
Delivery Location	_x_ Campus __On-line Distance Learning
Course Type	_x_ Standard (default) __Dissertation __Online Distance Learning __Other (specify: Placement, Student Led Individually Created Course, Year Abroad)
Marking Scheme By default, courses use a numerical marking scheme. If you wish to use a grade-only marking scheme, your course proposal below should justify this.	_x_ Standard (numerical) __Letter grade only

Guidance for remaining sections:

For an initial course proposal, please complete the **cover page and Section 1 (Case for Support)**, which asks you to describe the need for this course and to provide an overview of the course design, including the learning outcomes. **Please discuss your plans as early as possible with the head of Curriculum Review to avoid unnecessary effort.**

Send the form with these sections completed to the BoS Academic Secretary and head of Curriculum Review (listed on the BoS page) to obtain their comments before filling out the remainder of the form.

If a full proposal is invited, please complete the remaining sections and send to **iss-bos@inf.ed.ac.uk**.

2. Student-facing course description and additional feedback and assessment information.
This section provides most of the information students see in the DRPS entry for this course, as well as related details for BoS consideration.

3. Further information for BoS consideration: sample materials.

4. Additional Course Details required for DRPS. *[Administrative information such as delivery timing and prerequisites.]*

5. Placement in degree programme tables. *[Required for all level 9-11 courses; used to determine where the course will be added to existing degree programme tables.]*

6. Comments from colleagues. *[All course proposal should be sent to relevant colleagues in the area as well as to the appropriate year organizer and BoS Academic Secretary for comment in good time before the BoS meeting. Use this section to indicate what feedback has been solicited and received.]*

Colour coding and item-by-item guidance:

Guidance is provided in italics for each item. Please also refer to the guidance for new course proposals at http://www.inf.ed.ac.uk/student-services/committees/board-of-studies/course-proposal-guidelines. Examples of previous course proposal submissions are available on the past meetings page http://web.inf.ed.ac.uk/infweb/admin/committees/bos/meetings-directory but note that the proposal form was updated in Jan 2019.

Sections in gold are for student view and are required before a course can be entered into DRPS. You must complete these sections even if your course has already been approved based on other documentation.
Sections in orange are for School use but are still required for all courses (even those that have already been approved based on other documentation).
Section in gray are for consideration by the Board of Studies. They are normally required for all new course proposals but may be omitted in some circumstances (e.g., for invited course proposals) if you obtain permission in advance.

1. Case for support

This section is for consideration by the Board of Studies. The final two boxes (Learning Outcomes, Graduate Attributes) will also go into the student-facing course description.

Overall contribution to teaching portfolio and relation to existing curriculum

Please explain (a) what motivates the course proposal (e.g. a previous course having become outdated/inappropriate, an emergent or maturing research area or new research activity in the School, offerings of our competitors) and (b) how it relates to existing courses and degree programmes (including any prerequisite courses). Every new course should make an important contribution to the delivery of our Degree Programmes.

The course will introduce a mathematical development of machine learning to students who have already gained knowledge of machine learning in courses such as IAML or MLPR.

From the entry for DRPS:

Learning and inference algorithms are increasingly being used in many kinds of complex systems. In such applications, it is often important for the algorithm to have assurances of certain properties. Some of these properties are of social relevance, such as privacy (sensitive information is not revealed) and fairness (no bias against groups or individuals). Other properties such as stability, simplicity and confidence of learning models are important for reasons of robustness and interpretability in complex systems, for example, in autonomous systems.

The course will provide an introduction to formalising, interpreting and analysing such properties of machine learning algorithms.

Additional comments for BoS:

Analogous topics are taught at many top universities, including: Oxford, ETH, MIT, CMU, Cornell, Upenn, Warwick, Princeton, UMASS, U. Wisc-Madison, USC, NYU, Stanford and many others. The Caltech course by Yaser Abu-Mostafa is extremely popular in its online version, with millions of views.

The course will serve to satisfy some of the increasing demand for ML oriented courses at Informatics.

IAML or MLPR have been stated as recommended pre-requisite.

Target audience and expected demand

Describe the type of student the course would appeal to in terms of background, level of ability, and interests, and the expected class size for the course based on anticipated demand. A good justification would include some evidence, e.g. by referring to projects in an area, class sizes in similar courses, employer demand for the skills taught in the course, etc

The target of this course are mainly two groups of Informatics students:

1. Those who are interested in statistical AI and ML and would like to understand theoretical ideas related to the topic
2. Those who are interested in maths, theory and algorithms, and would like a mathematical introduction to machine learning (this may include some students from maths).

We expect 40 to 50 students in the course. For example, the analogous course at Oxford has about 45 students, out of only about 100 students eligible to take the course. Oxford has fewer competing ML courses, but also has fewer students interested in ML. Given that we have substantially larger number of possible students, with more of them interested in ML, we can expect similar to larger enrolments.

The course will run in Semster 2 so that students who take other ML courses such as IAML or MLPR and have a knowledge of ML and its applications can then take this course to get a theoretical understanding.

Anticipated Resource Requirements

Estimate how much lecturing, tutoring, exam preparation and marking effort will be needed in steady state, and any additional resources needed to set the course up initially. Provide estimates relative to class size where applicable and discuss how support staff will be recruited and supervised, if the class is likely to be very large. Please mention any scaling limits due to equipment or space. If equipment is required, say how it will be procured and maintained.]

We propose to have 3 tutorial sessions to help students get used to the mathematical material. These tutorials will be held around middle of the semester, possibly in a staggered schedule.

At steady state, we propose, in addition to the lecturer:

1 TA (small number of hours) to help update notes, tutorial materials, and answer student questions.

4 Tutors, 3 sessions each (based on anticipated class size)

1 marker to mark the coursework.

In the initial year a larger number of hours of a TA will be required to prepare the course tutorials and some notes. Lauren Watson, who is a PhD student with suitable background, is available for this task.

Quotas, special arrangements or unusual characteristics

Please specify if this course requires any special arrangements such as quotas or other registration arrangements; is a collaboration with another school or institution, or has other atypical characteristics that may affect finances or student registration. Further justification/information may be requested for such courses.

None

Narrative description of the course aims and structure

Please describe the main goals of the course and how the course design will allow students to achieve those goals. This section should be consistent with the student-facing information provided below, but should provide additional information to help colleagues at BoS understand the vision and structure of the course. This description may refer to the learning outcomes and graduate attributes (next two boxes) and should explain how activities such as tutorials, labs, or in-lecture activities will support them, and how the proposed assessments will assess them.

For courses that are important pre-requisites for other courses, this section may also provide content/syllabus information which is too detailed for the student-facing description, such as a lecture-by-lecture syllabus.

From the DRPS description:
The course will cover the following topics.

1. Characterising accuracy and confidence of learning models. E.g. probably approximately correct (PAC) guarantees
2. Complexity of learning models (e.g. VC dimension) and bias-complexity tradeoff
3. Learning low complexity models (Occam's razor): Structural risk minimisation, regularisation
4. Robustness properties of learning. E.g. stability, smoothness and Lipschitz properties
5. Kernel methods
6. Statistical notions of Privacy in learning. E.g. Differential privacy. And relation to complexity and robustness properties
7. Statistical/mathematical approach to Fairness. Individual and group fairness, and relation to privacy and other learning properties.

The topics will be discussed with reference to standard machine learning techniques, and examples of realistic problems. Our approach will be through precise mathematical definitions and analysis, with examples and counterexamples for illustration.

Tutorials and problem sets will be available to help improve understanding.

Additional comments for BoS:

The course will follow the textbook "Understanding Machine Learning: From Theory to Algorithms", by Shai Ben-David and Shai Shalev-Schwartz.

For topics such as privacy and fairness that are not covered in the book, reading materials and notes will be provided. In accordance to comments from several colleagues, the plan has been updated to increase emphasis on the privacy and fairness aspects. The current intention is to spend about 5 - 6 lectures on these topics. (Note that previous topics and mathematical formalisms will be introduced in a way to ease the coverage of these topics.)

The course will be assessed at 80% exam, 20% coursework.

Additional reference: The textbook "Foundations of Machine Learning" by Afshin Rostamizadeh, Ameet Talwalkar, and Mehryar Mohri.

Summary of Intended Learning Outcomes (MAXIMUM OF 5)

List the learning outcomes of the course. These must be assessable (i.e., observable), so must specify what the student should be able to do concretely, not simply what they should "understand". Use concrete verbs that indicate (a) what type of assessment would be appropriate, and (b) what level of knowledge/thinking is expected (from recall to analysis to novel creation). ***Example verbs:*** *define, explain, implement, compare, justify. Assessments (described later) should be tied to the learning outcomes.*

Outcomes should typically focus more on the types of thinking/skills developed than on the detailed course content, and the level of thinking should be appropriate to the level of the course: outcomes for a Level 11 course should include more higher-level thinking skills than for a Level 8 course. Further guidance on writing learning outcomes can be found at https://www.ncl.ac.uk/ltds/assets/documents/res-writinglearningoutcomes.pdf

On completion of this course, the student will be able to

1) Explain and interpret formal mathematical statements about properties of machine learning methods.

2) Critique and appraise properties of learning algorithms.

3) Compare and contrast the implications of privacy, fairness and other aspects of learning algorithms.

4) Formulate precise mathematical requirements to satisfy desired properties in real learning problems.

Graduate Attributes, Personal & Professional Skills

List the personal attributes and generic transferrable skills this course will help develop. Examples include

Cognitive skills: *problem-solving, critical/analytical thinking, handling ambiguity*

Responsibility, autonomy, effectiveness: *independent learning, self-awareness and reflection, creativity, decision-making, leadership, organization and time management, flexibility and change management, ethical/social/professional awareness and responsibility, entrepreneurship*

Communication: *interpersonal/teamwork skills, verbal and/or written communication, cross-cultural or cross-disciplinary communication*

Problem solving, critical/analytical thinking, independent learning, written communication.

1. Student-facing course description and additional feedback and assessment information

Except where noted, all fields are required and will go into the DRPS entry for the course (for use by students). Important: any text in DRPS is effectively a contract with students, so should not include details that are likely to change from year to year.

Summary Description *Provide a brief official description of the course, around 100 words. This should be worded in a student-friendly way, it is the part of the descriptor a student is most likely to read. If this course replaces another course, please say so in this summary.*	Learning and inference algorithms are increasingly being used in many kinds of complex systems. In such applications, it is often important for the algorithm to have assurances of certain properties. Some of these properties are of social relevance, such as privacy (sensitive information is not revealed) and fairness (no bias against groups or individuals). Other properties such as stability, simplicity and confidence of learning models are important for reasons of robustness and interpretability in complex systems, for example, in autonomous systems. The course will provide an introduction to formalising, interpreting and analysing such properties of machine learning algorithms.
Keywords *Give a list of searchable keywords.*	Machine learning, data science, algorithms, theory
Course Description 1. *A more detailed student-facing description of the course, which should normally include (a) a more in-depth academic description of the learning aims, nature and context of the course, (b) a rough outline of the content or syllabus, often as bullet points, and (c) a description of how the course will be taught, how students are expected to engage with their learning and how they will be expected to evidence and demonstrate their achievement of the intended learning outcomes*	The course will cover the following topics. 1. Characterising accuracy and confidence of learning models. E.g. probably approximately correct (PAC) guarantees 2. Complexity of learning models (e.g. VC dimension) and bias-complexity tradeoff 3. Learning low complexity models (Occam's razor): Structural risk minimisation, regularisation 4. Robustness properties of learning. E.g. stability, smoothness and Lipschitz properties 5. Kernel methods 6. Statistical notions of Privacy in learning. E.g. Differential privacy. And relation to complexity and robustness properties 7. Statistical/mathematical approach to Fairness. Individual and group fairness, and relation to privacy and other learning properties. The topics will be discussed with reference to standard machine learning techniques, and examples of realistic problems. Our approach will be through precise mathematical definitions and analysis, with examples and counterexamples for illustration.

	Tutorials and problem sets will be available to help improve understanding.
Assessment Weightings: *These should correspond approximately to the proportion of learning outcomes that each component assesses. More than 30% coursework requires specific justification.* *The expectation for a 10pt course is 20% coursework with the equivalent of one 15-20hr assessed assignment (but possibly split into smaller pieces). See 'components of assessment' below.*	Written Exam __80___% Practical Exam _____% (*for courses with programming exams)* Coursework __20___%
Further Assessment Information *Provide any further information that should go on DRPS for students. E.g., if the assessment includes required group work or if students must pass some individual component of assessment as well as the course overall.*	
Components of assessment and time spent on assignments (for BoS only) *If not already included in the course narrative description, please describe the type of assessments (oral presentation, report, programming, etc) and **how each component of assessment will assess the intended learning outcomes**. Where coursework involves group work, it is important to remember that every student has to be assessed individually for their contribution to any jointly produced piece of work.* *Also estimate **how many hours** students will spend on assignments. Please see the School policy on Workload and Assessment, which states that students should not be expected to spend more than 6-7 hrs/wk per 10 credits, including contact hours.* *Note that it often desirable to include formative assignments which are not formally assessed but submitted for feedback, often in combination with peer assessment.*	The coursework component is 20%. The coursework will be designed to take about 15 hours of time.
Feedback Information *Provide a high-level description of how and what type of feedback will be provided to students, for inclusion in DRPS.*	
Additional Feedback Information (for BoS use only) *If not already included in the course narrative, provide further details on planned feedback arrangements. This includes how course feedback is solicited from the class and responded to, as*	There will be a discussion session around the mid semester feedback. Additionally, a piazza will be available for students to discuss and ask questions.

<table>
<tr>
<td>well as what feedback students will get (either on work that contributes to their final mark, or not).

The University is committed to a baseline of principles regarding feedback that we have to implement at every level, and the School encourages submission of at least one piece of written work for formative feedback.

In general, formative feedback:
• Should say how students can improve.
• Need not be on individual work (e.g., consider a lecture or document summarizing common issues.)
• Can include oral feedback during labs/tutorials
• Can include feedback from peers
• Clickers/TopHat/equivalents can provide in-class feedback for both students and lecturer.
• Is returned in time for other forms of assessment to which it relates, to allow feedforward.</td>
<td>Additional notes and exercise problems will be provided based on student responses, common questions or confusions on piazza or tutorials.</td>
</tr>
<tr>
<td>Breakdown of Learning and Teaching Activities
State how many hours students spend on each part of the course. The total should be 10 x course credits, but please also see the School policy on Workload and Assessment.which states that students should not be expected to spend more than 6-7 hrs/wk per 10 credits, including contact hours.

Assume 10 weeks of lectures slots and 10 weeks of tutorials, but these need not all be used. As a guideline, a 10-pt course typically has 18-20 lecture hours, but should have only around 15 lectures of examinable material; the rest should be used for guest lectures, revision sessions, introductions to assignments, etc.</td>
<td>Contact hours
<table>
<tr><th>Hours</th><th>Type</th></tr>
<tr><td>18</td><td>Lecture Hours</td></tr>
<tr><td>3</td><td>Seminar/Tutorial Hours</td></tr>
<tr><td></td><td>Dissertation Project Supervision Hours</td></tr>
<tr><td></td><td>Supervised practical/Workshop/Studio hours</td></tr>
<tr><td>1</td><td>Feedback/Feedforward hours</td></tr>
<tr><td>2</td><td>Summative assessment hours</td></tr>
<tr><td>1</td><td>Revision Session Hours</td></tr>
</table>
Non-contact hours
<table>
<tr><th>Hours</th><th>Type</th></tr>
<tr><td>75</td><td>Directed Learning & Independent Learning hours</td></tr>
</table>
Total hours: 100</td>
</tr>
<tr>
<td>Reading List/Learning Resources
You are encouraged to create resource lists using LEGANTO</td>
<td>Book: “Understanding Machine Learning: From Theory to Algorithms”, by Shai Ben-David and Shai Shalev-Schwartz.</td>
</tr>
</table>

1. Further information for BoS consideration: sample materials

A full proposal for a new course must include examples of exercises and assessment. Please provide these below, along with publicity information if the course is to be advertised outwith the School.

Course information and publicity *The course web page (typically the Learn landing page) will be linked from the Sortable Course List, and information such as timetables and assignment deadlines must be made available prior to the start of the academic year. Please specify here if any additional info/publicity is needed for your course: typically only if it is aimed largely at non-SoI students.*	
Sample tutorial/lab sheet questions *Provide a list of tutorial questions and answers and/or samples of lab sheets. These need not be fully fleshed out but should indicate what sort of exercises will be provided to help students learn the material.*	
Sample assessment materials *If the course is primarily assessed by **exam**, provide a sample exam question with model answers. Any non-standard exam format must be justified. The online list of past exam papers gives an idea of typical and alternative exam formats: http://www.inf.ed.ac.uk/teaching/exam_papers/.* *If the course is largely or primarily assessed by **coursework**, provide a sketch of a possible assignment with an estimate of effort against each sub-task and a description of marking criteria.*	1. Define kernels. Suppose there is a classification problem, where there is a cluster of blue points within the unit disk around origin o and another cluster of red points within the annulus $4 < \|p - o\| < 5$. Define a polynomial kernel that can be used for this classification. What is the lifting map \phi for your Kernel? Describe a linear separator for the classes in the higher dimensional space. Ans: Kernels are measures of similarity defined on a set X of elements. Formally, they are functions K:XxX -> H for a Hilbert space H, such that corresponding to K there is a \phi : X-> H which satisfies K(u, v) = <\phi(u), \phi(v)>, where <,> is an inner product in H. For the Given problem, a suitable Kernel is $K(u,v) = (u.v)^2$. The corresponding lifting map is given by $\phi(u) = (u_x^2, u_y^2, \sqrt{2}\, u_x u_y)$. A linear separator is given by Z = 2, or the plane of points (*, *, 2). 2. Define Replace-one-average-stability. Show with an example that such a stability does not imply small loss. Show with an example that it does not imply privacy in the sense that an adversary may be able to deduce the presence or absence of some point in the training sample based on the learning output. [Use as example a learning algorithm outputs a model with linear decision boundary and is replace-one-average-stable.] Ans: An algorithm A is replace-one-average-stable if there is a monotone non-increasing function f:N -> R such that the expected loss: $E[L(A(S^i))-L(A(S))] < f(m)$ for sample S of size m, and S^i represents S with one point i replaced or removed. For the example: Suppose all points are in the square [0,1]x[0,1]. And suppose positive points are exactly those with x+y < 1. So the true decision boundary is x+y=1. Now suppose A outputs boundary x = 0.5 on all training samples. This implies stability, but high loss. On the same example, suppose there is a particular point i, on whose removal (i.e. for input set S^i), A outputs decision boundary y = 0.5. This model has the same loss as before, but the output model of A is different on S^i, and therefore and adversary can deduce the presence/absence of i. (Implying that this is not differentially private.) 3. Suppose an IoT controlled light "learns" with use when to turn on and off. If the learning algorithm lacks stability, state some negative effects that may have. Ans: Lack of stability will mean that addition of new data points can cause large changes to its behaviour. Thus, as it learns: * It may change its daily schedule dramatically day to day

	* If the input consists of behaviour/activity information of the user (e.g. data from motion sensors) then different actions may set it off on different days leading to further frustrations.
Any other relevant materials *Include anything else that is relevant, possibly in the form of links. If you do not want to specify a set of concrete readings for the official course descriptor, please list examples here.*	

2. Additional Course Details for DRPS

Except where otherwise noted, these fields are required for entry into EUCLID and will be visible to students in the DRPS entry.

Planned Academic Year of Delivery *(The first year you anticipate the course running, e.g. AY 2019-20)*	
Course Organiser *(By default, the course proposer)*	
Intended Delivery Period	___Semester 1 __x_Semester 2 ___Full Year ___Summer ___Other (please specify):
Timetable considerations/conflicts *For School use. Please specify any constraints to be considered (e.g. overlap of popular combinations, other specialism courses, external courses etc). Include whether the semester delivery is constrained or could be flexible.*	
Is this course available to visiting students?	___Yes (default) ___No **If no, please provide a justification here:**
Required pre-requisite courses *Use sparingly: these are enforced in PATH and can only be waived by approval from the School's Curriculum Approval Officer. Note that cross-year required pre-requisites may prevent MSc students from registering; consider using recommended pre-requisites or "other requirements" instead.*	___No ___Yes (please specify full course name(s) and code(s)):
Recommended pre-requisite courses	___No __x_Yes (please specify full course name(s) and code(s)): INFR10069: Introductory Applied Machine Learning OR INFR11130: Machine Learning and Pattern Recognition

Required co-requisite courses *Specify any courses that must be taken in parallel with the existing course. Note that this leads to a timetabling constraint that should be mentioned elsewhere in the proposal.*	___No ___Yes (please specify full course name(s) and code(s)):
Prohibited Combinations *Specify any courses that may not be taken in combination with the proposed course].*	___No ___Yes (please specify full course name(s) and code(s)):
Other Requirements/Additional Information *This information is often used by MSc students and students from other Schools to see if they have appropriate background without having done our School's courses. So please avoid course titles, instead list specific knowledge and skills (such as mathematical concepts, programming ability or specific languages, etc).* *Also list any other constraints on registration, for example: "Only available to 4th Year Informatics students including those on joint degrees." or "This course is open to all Informatics students including those on joint degrees, and to students in the School of Mathematics. Other external students whose DPT does not list this course should seek permission from the course organiser."*	___No _x__Yes (please specify): Students should be confident in standard machine learning ideas: training and test sets, classification, regression, clustering; standard machine learning methods: support vector machines, linear regression, k-means etc. Good understanding of probability and probabilistic arguments is necessary.
Visiting Student Pre-requisites	___Same as "other requirements" ___Different than "other requirements" (please specify):

3. Placement in degree programme tables: for level 9-11 courses only

This section is for consideration by the Board of Studies and will be used later by ITO to determine where the course will be added to existing degree programme tables.

Is this course restricted to students on a specific degree? *E.g., some courses are only available to students on a specific CDT or MSc.*	_x_ No ___ Yes (please specify and provide justification):
Is this course compulsory for students on any degree(s)?	_x_ No ___ Yes (please specify and provide justification):
Any issues for part-time students? *Normally, part-time students have access to the same courses as full-time students on the equivalent degree. If you anticipate any problems with this, please specify here.*	No.

For optional courses:

If this course is available but non-compulsory for students on various degrees (most courses), please fill in this section. The choices here determine where the course appears in degree programme tables (DPTs) and the 2-3 character tags are displayed in the Informatics sortable course list.

Should this course be tagged as 'ML' (machine learning foundations and methods)? *Courses with the ML tag are typically very high-demand and most degrees limit the number of ML credits. If your course might appeal to a similar audience but draw off students from these large courses, please select 'no' and choose one of the tags below.*	__ No __ Yes
If you chose 'no', please choose at least one of the following tags… *Ideally, select exactly one, unless there is a good argument for more than one. These three are used in various combinations for many of our degrees.*	_x_ **FSS** (CS foundations, systems, and software) _x_ **AIA** (artificial intelligence applications and paradigms) __ **COG** (cognitive science: including HCI and NLP courses, but not most other AI courses. Please restrict to courses most relevant to natural cognition.)
…and also tick if any of the following tags or categories apply. *Do not tick any of these if you selected 'ML' already.*	__ **NS** (natural systems: e.g., computation by or about biological or social systems. Many COG courses are also NS. This tag is mainly relevant for MSc in Informatics.) __ **SE** (software engineering: including courses that are highly relevant to SE degrees. All SE courses should also be FSS. This tag is mainly relevant for UG SE degrees.) __ Databases and data management systems (used for Data Science MSc and MSc(R))

	_x_Unstructured data and applications (used for Data Science MSc and MSc(R)) __Level 11 Security courses (used for Security MSc)
If you are not sure which tags are most appropriate or have other questions about this section, please note any comments/issues here.	

4. Comments from colleagues

All course proposal should be sent to relevant colleagues in the area as well as to the appropriate year organizer and BoS Academic Secretary for comment in good time before the BoS meeting. Please indicate here what feedback has been solicited and received.

Additional Comments *Summarise any comments received from relevant individuals prior to proposing the course. If you have not discussed this proposal with others please note this.*	Several colleagues have been consulted. And comments integrated into the proposal. Ram Ramamoorthy: This is a necessary course. But clarify fit with existing courses. [Done.] Michael Gutmann, Iain Murray: The description mentions some ML algorithms covered elsewhere (e.g. IAML). [Clarified that this is supposed to touch on some of those topics from the perspective of the theoretical analysis.] Iain Murray: List of topics is long, and needs better explanation. [List of topics adjusted to be more precise, and explanations provided for better interpretation by students.] May be good to focus it a bit with the aim to build toward something of current interest. This is best done during course material preparation while keeping the DRPS general enough to cover this possibility. [Agreed. The current sequence naturally leads up to privacy and fairness discussions. To be improved during course prep.] Amos Storkey: Positive about the course. Suggest slight adjustment to the name. "Theoretical *Foundations.*." may cause students to misinterpret since ML theory is a broader subject. Other possible names: * Computational Learning Theory * Learning theory * Theoretical Analyses in Machine Learning.. [Note that some other materials (outside of conventional learning theory) like those mentioned in Amos' comment in the content description are to be covered.] Paul Patras: "As machine learning is increasingly adopted throughout many critical services, practice based experience will not be enough. Students should be able think critically about why certain models behave in a certain way, have confidence in their output, ensure fairness, safety and explainability. This course seems to provide an introduction to such rigorous theoretical reasoning about learning." Other suggestions: * Clarify that fairness etc is not be just 1 lecture [done.] * May need more TAs [Possibly, but not a lot more. To be considered.] * Think about precise pre-requisites and hours distribution. [To do.] * Example questions would be good. [Done.] Amos storkey: There are other ways to characterise ML algorithms, such as regret, risk measures, convergence, learning curves, Bayesian theory etc. These topics will be considered for inclusion in future years. The DoT recommendation is that in the first year it is best to follow a simpler well defined plan.

Year Organiser Comments *Year Organisers are responsible for maintaining the official Year Guides for every year of study, which, among other things, provide guidance on available course choices and specialist areas. The Year Organisers of all years for which the course will be offered should be consulted on the appropriateness and relevance on the course. Issues to consider here include balance of course offerings across semesters, subject areas, and credit levels, timetabling implications, fit into the administrative structures used in delivering that year.]*	UG4, 5, MSc year organisers have been consulted. Walid Magdy: Clarify relation to other ML courses. [Done].
BoS Academic Secretary Comments *Proposals must be checked by the Secretary of the Board of Studies prior to discussion at the actual Board meeting. This is a placeholder for their comments, mainly on the formal quality of the content provided above.*	