

COURSE OUTLINE: MH3520

Course Title	Mathematics of Deep Learning		
Course Code	MH3520		
Offered	Study Year 3, Semester 1		
Course Coordinator	Philipp Harms (Dr)	philipp.harms@ntu.edu.sg	6513 7187
Pre-requisites	MH2100 and MH3500 and MH3600 and PS0001		
AU	4		
Contact hours	Lectures: 39, Tutorials: 12		
Approved for delivery from	AY 2022/23 semester 1		
Last revised	8 Jul 2022, 18:57		

Course Aims

This course investigates deep learning from the perspectives of several mathematical theories: numerical optimisation, statistical learning, function approximation, and coding theory. The aim is to shed some light on why and under what circumstances deep learning can be expected to work well - or not.

Intended Learning Outcomes

Upon successfully completing this course, you should be able to:

1. Describe numerical optimization techniques in deep learning
2. Formulate deep learning as a statistical learning problem for a quantitative description of the inherent complexity
3. Use function approximation and coding theory to derive upper and lower bounds on the approximation power of neural networks
4. Implement deep learning algorithms in Python

Course Content

History of deep learning and relation to other machine learning techniques

Training of neural networks via numerical optimization algorithms such as gradient descent

Error bounds for deep learning algorithms via statistical learning theory

Universality of neural networks via the Stone-Weierstrass theorem

Approximation rates for neural networks via polynomial/spline/wavelet approximation and coding theory

Assessment

Component	Course ILOs tested	SPMS-MAS Graduate Attributes tested	Weighting	Team / Individual	Assessment Rubrics
Continuous Assessment					
Tutorials					
Homework	1, 2, 3, 4	1. a, b, c, d 2. a, b 3. a 4. a	20	individual	See Appendix for rubric
Presentation	1, 2, 3, 4	1. a, b, c, d 2. a, b 3. a 4. a	20	individual	See Appendix for rubric
Examination (2 hours)					
Short Answer Questions	1, 2, 3	1. a, b, c 2. a, b 3. a 4. a	60	individual	See Appendix for rubric
Total			100%		

These are the relevant SPMS-MAS Graduate Attributes.

1. Competence

- a. Independently process and interpret mathematical theories and methodologies, and apply them to solve problems
- b. Formulate mathematical statements precisely using rigorous mathematical language
- c. Discover patterns by abstraction from examples
- d. Use computer technology to solve problems, and to communicate mathematical ideas

2. Creativity

- a. Critically assess the applicability of mathematical tools in the workplace
- b. Build on the connection between subfields of mathematics to tackle new problems

3. Communication

- a. Present mathematics ideas logically and coherently at the appropriate level for the intended audience

4. Civic-mindedness

- a. Develop and communicate mathematical ideas and concepts relevant in everyday life for the benefits of society

Formative Feedback

The proportion of your solved homework is continuous feedback on your ability to follow the class. Moreover, in the tutorial sessions, you will receive individual feedback when you present your solutions and collective feedback when we discuss common mistakes.

Learning and Teaching Approach

Lectures (39 hours)	You will be introduced to deep learning from the perspectives of several mathematical theories. There will be excursions to fill you in on mathematical background as needed, based on informal feedback sessions.
Tutorials (12 hours)	In your homework, you will reinforce mathematical concepts covered in the lectures, learn how to implement deep-learning algorithms in Python, and interpret the numerical evidence in the light of theoretical predictions. When presenting your solutions, you will train to communicate mathematical ideas and arguments.

Reading and References

There is no single textbook covering the content of the course, and we will have to glimpse into various pieces of original literature instead. That being said, the following lecture notes can serve as a guide for the developments in the course:

Philipp C. Petersen (University of Vienna): Neural Network Theory. pc-petersen.eu/Neural_Network_Theory.pdf

Course Policies and Student Responsibilities

(1) Attendance

You are expected to actively participate in all lectures and tutorials. If you miss a lecture, you must inform the instructor via email prior to the start of the class. If you miss a tutorial, you must obtain official leave from the administration. Unexcused absence will negatively impact your grade.

(2) Homework

Before each tutorial, you will be asked to indicate which homework problems you have solved. For any problem marked as solved, you must be able to present a complete and rigorous solution when requested to do so.

Academic Integrity

Good academic work depends on honesty and ethical behaviour. The quality of your work as a student relies on adhering to the principles of academic integrity and to the NTU Honour Code, a set of values shared by the whole university community. Truth, Trust and Justice are at the core of NTU's shared values.

As a student, it is important that you recognize your responsibilities in understanding and applying the principles of academic integrity in all the work you do at NTU. Not knowing what is involved in maintaining academic integrity does not excuse academic dishonesty. You need to actively equip yourself with strategies to avoid all forms of academic dishonesty, including plagiarism, academic fraud, collusion and cheating. If you are uncertain of the definitions of any of these terms, you should go to the [Academic Integrity website](#) for more information. Consult your instructor(s) if you need any clarification about the requirements of academic integrity in the course.

Course Instructors

Instructor	Office Location	Phone	Email
Philipp Harms (Dr)	SPMS-MAS-05-41	6513 7187	philipp.harms@ntu.edu.sg

Planned Weekly Schedule

Week	Topic	Course ILO	Readings/ Activities
1	Introduction to Deep Learning	1, 4	Lecture
2	Basics of Numerical Optimization	1, 4	Lecture and Tutorial
3	Training Neural Networks	1, 4	Lecture and Tutorial
4	Basics of Probability Theory	2, 4	Lecture and Tutorial
5	Error Bounds for Deep Learning via Statistical Learning Theory	2, 4	Lecture and Tutorial
6	Basics of Functional Analysis	3, 4	Lecture and Tutorial
7	Universality of Neural Networks	3, 4	Lecture and Tutorial
8	Basics of Approximation Theory	3, 4	Lecture and Tutorial
9	Kolmogorov-Arnold Representation and Approximation by Networks of Bounded Size	3, 4	Lecture and Tutorial
10	Splines and Approximation by Wide Networks	3, 4	Lecture and Tutorial
11	Wavelets and Approximation by Wide Networks	3, 4	Lecture and Tutorial
12	Coding Theory and Best Approximation Rates	3, 4	Lecture and Tutorial
13	Analyticity and Approximation by Deep Networks	3, 4	Lecture and Tutorial

Appendix 1: Assessment Rubrics

Rubric for Tutorials: Homework (20%)

Point-based marking

Rubric for Tutorials: Presentation (20%)

For any problem indicated as solved, you must be able to present a complete and rigorous solution when requested to do so in the tutorial. You will be requested at least 2 or 3 times (depending on class size) to present a solution and will be scored on it individually as follows.

Grading Criteria	High standard (20 – 14)	Average standard (13 – 7)	Low standard (0 – 6)
Quality of the solution	The solution is accurate and complete.	The solution is mostly accurate and complete, up to some minor mistakes or missing pieces.	There is a major mistake or missing piece.
Quality of the presentation	The presentation is clear and well organized. It is easy to follow your arguments.	The presentation is mostly clear and organized. It is possible to follow your arguments.	The presentation is unclear or messy. It is difficult or impossible to follow your argument.

Rubric for Examination: Short Answer Questions (60%)

Point-based marking