COURSE MANUAL
DEEP LEARNING

Research Master Business Data Science
Erasmus University Rotterdam, University of Amsterdam, Vrije Universiteit Amsterdam

Course Code	EBDS19107
Academic year	2019-2020
Period	1.5
Credits	4
Recommended knowledge	Machine Learning
Required knowledge	Linear algebra, Regression,

Table of Contents

1. COURSE COORDINATOR AND LECTURERS

Coordinator/Lecturer: Dr. Eran Raviv (APG-AM)
Email: eeraviv@gmail.com

Short bio: Eran Raviv holds a PhD in econometrics from Tinbergen Institute, a master's degree in applied statistics from Tel Aviv University and a second master's degree in quantitative finance from Rotterdam University. His research has been published in high impact peer-reviewed journals. In 2013 he joined APG-AM as a senior quantitative investment strategist. Since spending two years in that role, he has served as a data scientist working on various innovative projects across APG-AM organization, focusing on deep learning methodologies and NLP applications.

Teaching Assistant: TBD
Email: TBD

2. COURSE CONTENT

Deep learning course covers theoretical and practical aspects, state-of-the-art deep learning architectures, and application examples.

Topics covered:

1. Introduction to Deep Learning (High-level definitions of fundamental concepts and first examples)
2. Deep Learning components (gradient descent models, loss functions, avoiding over-fitting, introducing asymmetry)
3. Feed forward neural networks
4. Convolutional neural networks
5. Embeddings (pre-trained embeddings, examples of pre-trained models, e.g., GloVe embeddings, Word2Vec)
6. Recurrent neural networks
7. Long-short term memory units
8. Advanced architectures (Densely connected networks, Adaptive structural learning)

3. LEARNING OBJECTIVES

By the end of the course students will be able to:

- Understand the fundamental building blocks of deep learning methods,
- understand the weaknesses and strengths of the different architectures,
- know how to tackle weaknesses and tailor the model for a particular application,
- program these methods, and
- be able to describe the numerical computational steps applied by the machine.

KNOWLEDGE AND UNDERSTANDING	Understand the fundamental building blocks of deep learning methods, Introduction to Deep Learning (theory and practice) - Introduction to Deep Learning (High-level definitions of fundamental concepts and first examples) - Deep Learning components (gradient descent models, loss functions, avoiding over-fitting, introducing asymmetry) - Feed forward neural networks - Convolutional neural networks - Embeddings (pre-trained embeddings, examples of pre-trained models, e.g., GloVe embeddings, Word2Vec) - Recurrent neural networks - Long-short term memory units - Advanced architectures (Densely connected networks, Adaptive structural learning)
APPLICATION OF KNOWLEDGE	- ability to program the methods - develop familiarity with the broad range of applications (text, audio, visual)

MAKING JUDGEMENT	- ability to tailor the models for particular application - ability to tackle weaknesses of each methods
COMMUNICATION	- ability to communicate the architecture of the model - ability to communicate the internal numerical optimizations
LEARNING SKILLS	- ability to keep abreast with ongoing research in the field

4. STUDY MATERIAL

The following list of mandatory readings (presented in alphabetical order) is considered essential for your learning experience. These articles are also part of the exam material. Changes in the reading list will be communicated on CANVAS.

Books:

- *Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press.*
- *Patterson, J. and Gibson, A., 2017. Deep learning: A practitioner's approach.*

Selected papers, including:

- Frank Z Xing, Erik Cambria, and Roy E Welsch. Natural language based financial forecasting: a survey. Artificial Intelligence Review, 50(1):49–73, 2018
- Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4700–4708, 2017
- Heaton, J.B., Polson, N.G. and Witte, J.H., 2017. Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), pp.3-12.
- Honglak Lee, Peter Pham, Yan Largman, and Andrew Y Ng. Unsupervised feature learning for audio classification using convolutional deep belief networks. In Advances in neural information processing systems, pages 1096–1104, 2009.
- Omer Levy and Yoav Goldberg. Neural word embedding as implicit matrix factorization. In Advances in neural information processing systems, pages 2177–2185, 2014
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), pp.1929-1958.

Lecture notes available on CANVAS.

5. FORM OF TUITION

The lectures aim at stimulating your academic skills, and providing you with new knowledge. In this course, lectures are accompanied by tutorials and computer lab sessions. The information provided in the lecture is essential for the assignments and discussions during those sessions. We expect students to come to the lectures well prepared and to participate in the interaction.

The tutorials aim at practicing the theory using exercises and allowing students to ask for additional explanation for those parts of the material perceived as more difficult.

The computer lab sessions aim at making the material come alive and train students in how the methods learnt in class can actually be applied to data. The lab sessions are meant to work on the assignments, such that you automatically keep up with the material.

Class participation is a part of the course that is important, asking questions, contributing to answers and to the general learning atmosphere.

6. ASSESSMENT

Your overall course grade is composed based on different components that are presented in the assessment overview. You need a minimum overall grade of 5.50 to pass the course.

The final grade is rounded to the nearest multiple of .0 or .5, with the following exceptions: any grade between 5.0 and 5.5 is rounded to a 5; a 5.5 is rounded to a 6; a 0.5 does not exist. Grades for homework or midterm examinations do not need to be rounded.

Format	% grade	Knowledge and Understanding	Application of knowledge	Communication	Learning Skills
Exam	85%	X			
Assignments	15%*	X	X	X	X *Assignments are weighted by the student individual participation to the team.

Group Assignment - group assessment
During the first tutorial session, you will be assigned to a group of 3 students. The group assignments are designed to challenge you in various ways. They bring together different materials you study and practice during the lectures and tutorials, illustrating (some of) the principles learned in the course work and their applications (see **Appendix A**).

Group Participation - individual assessment
To assess your participation in the team we use an online peer evaluation system. You will assess both your teammates and yourself using a questionnaire (see **Appendix B**). This will result in a weighting factor that reflects the amount of effort and skills you have put in the group assignment. It consists of the average of the scores that the others give you on your role in the team and the assignment. From the early start of the project you know upon which social skills you and your peers will evaluate each other. This means you can pay extra attention to these competences. These are the categories that you will focus on when evaluating your team members and yourself: Contributing to Work; Interacting with Teammates; Keeping Team on Track; Working Quality; Having Knowledge/Skills; and Team Satisfaction.

Class Participation - individual assessment
During the tutorials, you will be required to discuss, present, exchange information with fellow students and debate in different group settings. Hence, your presence is valued and essential in order for you to accomplish the objectives of this course (see **Appendix C** for the Assessment criteria).

Written exam - individual assessment

The final exam takes place in the last week of the course. The actual exam time is 2 hours in total. It will involve open-ended questions, which are testing knowledge, insight and application. During the last week of the course, we will practice a few exam questions. The open-ended question templates are provided in **Appendix D**. At the end of every session, a reference list of the compulsory papers that may be part of the exam will be shared in class.

7. DETAILED COURSE SCHEDULE

Please check Canvas for an up-to-date schedule, reading material and assignments.

Week	Date	Time	Format	Theme/Topics	Preparation
1			Lecture 1 (3hr)	**Introduction to Deep Learning &Deep Learning components** (gradient descent models, loss functions, avoiding over-fitting, introducing asymmetry)	Mandatory readings: *Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press. Ch 5.*
			Tutorial 1 (1hr)	**Setting the Scene**	Managing expectations & Project Kick-Off • Team formation • Explain assignment 1 • Explain what we expect
2			Lecture 2	**Neural networks**	*Mandatory readings:* • *Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press. Ch 6.*
			Tutorial 2 (1hr)		
			Submission Deadline	**Presentation**	Submit your presentation before 08.00 via canvas (see section 9.2)
3			Lecture 3	**Convolutional neural networks**	*Mandatory readings:* • *Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press. Ch 9.*
			Tutorial 3 (1hr)		• Presentations assignment 1 • Explain assignment 2 (see section 9.3)
			Submission Deadline	Assignment 1	Submit your assignment before 17.00 via canvas
4			Lecture 4	**Embeddings**	*Mandatory readings:* • Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.
			Tutorial 4 (1 hr)	**Q & A**	Team appointment (see section 9.4)
			Submission Deadline	**Presentation**	Submit your presentation before 08.00 via canvas (see section 9.2)

Week	Date	Time	Format	Theme/Topics	Preparation
5			Lecture 5	**Recurrent neural networks and LSTM**	*Mandatory readings:*
			Tutorial 5	**Presentation**	Presentation assignment 2
			Submission Deadline	Assignment 2	Submit your assignment before 17.00 via canvas
6			Lecture 6	**Advanced architectures and applications**	*Mandatory readings:* • Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition.
			Tutorial 6	**Exam Q&A**	Exam preparation (see section 9.5)
7			Individual study and team work		
			Exam Submission Deadline	**Presentation**	Submit your presentation before 08.00 via canvas (see section 9.6)
	TBA		**Exam inspection**	**Feedback**	Check Canvas for updates

8. ASSIGNMENT INFORMATION

9.1 Group assignment 1
In this assignment, you will choose your own dataset to work own, from a list of available public data. You will be required to code a very basic network and use it for prediction. It can be of the regression form, or classification form. An example could be: what should be a house price given a number of important characteristic (square meters, room numbers, location etc.). The second part of the first assignment is to describe the data and findings. Results should be reported in a concise manner in essay format (1- 2 pages, code and figures excluded).

9.2 Presentation assignment 1
You need to present your findings of assignment 1. Make a presentation for maximally 10 minutes, and prepare your presentation as a pitch. Thus, do not explain everything but highlight your most interesting insights. Explain the process and your main conclusions. Usually, presenting one slide requires 2-3 minutes.

9.3 Group assignment 2
In this assignment you will choose a paper to be read and presented. Your aim will be to understand the main motivation, methods used and conclusion of the paper. You are free to choose any deep learning related paper, as long as it was published in a reputable book or academic journal. The output should be a 10 minutes presentation of the paper chosen.

9.4 Team Q&A

Prepare yourself with questions you would like to ask the teacher. Per team a limited amount of time can be reserved so focus on the most important issues.

9.5 Exam Q&A
Prepare yourself with questions you would like to ask the teacher/TA. Focus on the most important issues.

9.6 Final presentation
In this final presentation, you present your two assignments, and discuss how they are connected. It is especially important to discuss progress in the accumulation of knowledge. Explain what would you have done differently in assignment one given the knowledge you gathered from assignment two.

business data science
GRADUATE PROGRAM | Erasmus VU

APPENDIX A –ASSIGNMENTS ASSESSMENT A

Criterion	5 or lower	6	7	8	9 or 10
Knowledge/Application of knowledge **Weight 50%**	The assignment does not address the question(s). It contains evident logical errors or omissions. The answer is too simple or too limited for program or study load.	The assignment addresses sufficiently the question(s). It does not reflect all material covered, it includes considerable simplifications or shortcuts. Minimum level of adequacy for study load.	The assignment clearly addresses the question(s). It shows good knowledge of the material covered, awareness and own reflection of the important aspects of the material covered.	Well-considered and well-explained answer. Clear evidence of very good knowledge and understanding of the material covered, without being exceptional.	The assignment addresses fully the question(s). The students fully master the material covered in the course.
Application of knowledge **Weight 20%**	Insufficient ability to apply the knowledge covered. Major flaws.	Sufficient ability to apply the knowledge covered. Despite several flaws in the assignment, the outcome is satisfactory.	Good ability to apply the knowledge covered, with a few oversights.	Very good ability to apply the knowledge covered in the course.	Excellent ability to apply the knowledge covered in the program to different settings/data.
Written communication **Weight 5%**	Unstructured text. Fails to convey the key message of the thesis and/or to address questions. The text does not meet the academic editorial standards.	The text is somewhat unstructured and unclear. The text barely passes the academic editorial standards, as more polishing work is needed.	Overall well written, with occasional typos, or inaccuracies. The text passes the academic editorial standards, although the writing style is mechanical.	Structured text. The text is clear and concise, but here and there more (or less) details could improve the readability. Tables and Figures are self-explicatory and timely introduced in the text. The text meets the academic editorial standards, although the writing style is a bit mechanical at times.	Structured, coherent and polished text. Excellent writing style. The text is accurate, clear and concise, with the right level of detail. Tables and Figures are self-explicatory and timely introduced in the text. The text meets the academic editorial standards.

Presentation **Weight 5%**	Unstructured presentation. Fails to convey the key message of the thesis and/or to address questions.	Structured oral presentation. Answers and comments by the audience are not always adequately addressed.	Structured oral presentation. Answers and comments by the audience are adequately addressed most of the times.	Good oral presentation: well structured, right level of detail. Gives accurate and to the point response to comments and questions	Excellent oral presentation: coherent, right level of detail, lively. Gives accurate and to the point response to comments and questions
TOTAL					

APPENDIX B – GROUP PARTICIPATION ASSESSMENT

Criterion	**Poor**	**Fair**	**Good**	**Very Good**	**Excellent**
Contributing to the team's work **Weight 20%**	Does not do a fair share of the team's work. Delivers sloppy or incomplete work. Misses deadlines. Is late, unprepared, or absent for team meetings. Does not assist teammates. Quits if the work becomes difficult.	Demonstrates behaviors described immediately left and right.	Completes a fair share of the team's work with acceptable quality. Keeps commitments and completes assignments on time. Helps teammates who are having difficulty when it is easy or important.	Demonstrates behaviors described immediately left and right.	Does more or higher-quality work than expected. Makes important contributions that improve the team's work. Helps teammates who are having difficulty completing their work.
Interacting with teammates **Weight 20%**	Interrupts, ignores, bosses, or makes fun of teammates. Takes actions that affect teammates without their input. Does not share information. Complains, makes excuses, or does not interact with teammates. Is defensive. Will not accept help or advice from teammates.	Demonstrates behaviors described immediately left and right.	Listens to teammates and respects their contributions. Communicates clearly. Shares information with teammates. Participates fully in team activities. Respects and responds to feedback from teammates.	Demonstrates behaviours described immediately left and right.	Asks for and shows an interest in teammates' ideas and contributions. Makes sure teammates stay informed and understand each other. Provides encouragement or enthusiasm to the team. Asks teammates for feedback and uses their suggestions to improve.
Keeping the team on track **Weight 20%**	Is unaware of whether the team is meeting its goals. Does not pay attention to teammates' progress. Avoids discussing team problems, even when they are obvious.	Demonstrates behaviors described immediately left and right.	Notices changes that influence the team's success. Knows what everyone on the team should be doing and notices problems. Alerts teammates or suggests solutions when the team's success is threatened.	Demonstrates behaviors described immediately left and right.	Watches conditions affecting the team and monitors the team's progress. Makes sure that teammates are making appropriate progress. Gives teammates specific, timely, and constructive feedback.
Expected quality	Satisfied even if the team does not meet assigned standards. Wants the team to avoid work, even if it hurts the team. Doubts that the team can meet its requirements.	Demonstrates behaviours described	Encourages the team to do good work that meets all requirements. Wants the team to perform well enough to earn all available rewards. Believes	Demonstrates behaviours described	Motivates the team to do excellent work. Cares that the team does outstanding work, even if there is no additional reward. Believes that the team can do excellent work.

		immediately left and right.	that the team can fully meet its responsibilities.	immediately left and right.	
Weight 20%					
Having related knowledge, skills, and abilities **Weight 20%**	Missing basic qualifications needed to be a member of the team. Unable or unwilling to develop knowledge or skills to contribute to the team. Unable to perform any of the duties of other team members.	Demonstrates behaviours described immediately left and right.	Demonstrates sufficient knowledge, skills, and abilities to contribute to the team's work. Acquires knowledge or skills as needed to meet requirements. Able to perform some of the tasks normally done by other team members.	Demonstrates behaviors described immediately left and right.	Demonstrates the knowledge, skills, and abilities to do excellent work. Acquires new knowledge or skills to improve the team's performance. Able to perform the role of any team member if necessary.
TOTAL					

business
data science
GRADUATE PROGRAM | Erasmus VU

APPENDIX C – CLASS PARTICIPATION ASSESSMENT

Criterion	Poor	Average	Good	Excellent
Level of engagement	Student never contributes to class by offering ideas and asking questions and/or has trouble staying on task during group project time.	Student rarely contributes to class by offering ideas and asking questions and/or works on group project only some of the allotted time.	Student proactively contributes to class by offering ideas and/or asks questions once per class and/or works on group project for most of the allotted time.	Student proactively contributes to class by offering ideas and/or asks questions more than once per class and/or works consistently on group project the entire time.
Weight 33%				
Quality of comments	Comments are uninformative, lacking in appropriate terminology. Heavy reliance on opinion and personal taste, e.g., “I love it”, “I hate it”, “It’s bad” etc.	Comments are sometimes constructive, with occasional signs of insight. Student does not use appropriate terminology; comments not always relevant to the discussion.	Comments mostly insightful and constructive; mostly uses appropriate terminology. Occasionally comments are too general or not relevant to the discussion.	Comments always insightful and constructive; uses appropriate terminology. Comments balanced between general impressions, opinions and specific, thoughtful criticisms or contributions.
Weight 33%				
Listening skills	Student does not listen when others talk, both in groups and in class. Student often interrupts when others speak. Student displays disruptive behavior during class.	Student does not really listen when others talk, both in groups and in class.	Student listens when others talk, both in groups and in class.	Student listens when others talk, both in groups and in class. Student incorporates or builds off of the ideas of others.
Weight 33%				
TOTAL				

APPENDIX D – EXAMPLES OF EXAM QUESTIONS

Question 1

Any deep learning model uses an activation function.

a. Why not use a linear activation function? [The gradient is constant and does not depend on the input]
b. Name one issue with using a sigmoid activation function. [Due to the problem of "vanishing gradients"]
c. Describe the problem of "vanishing gradients". Which activation function works around the problem? [Wide regions of the inputs has very small gradient, which makes the network extremely slow to train. ReLu works around the problem by having a constant gradient for wide range of inputs, while at the same time it less computationally expensive.]

Question 2

Any deep learning model has hyper parameters which need to be specified. Give an example for such hyper parameters, and explain couple of possible ways in which it can be tuned are. [Learning rate, number of layers, and activation function are all hyper parameters. One way to tune is to simply trial and error. Another way is to use a framework for structured testing of the hyper parameters and advise on the best one (e.g. Adanet)].

Question 3

a. Name and describe 3 possible applications of deep learning. [Machine translation, Machine recognition, word embedding, audio generation]
b. Name 3 open source tools which facilitate the construction of deep learning models [Tensorflow, Keras, Pytoch, H2O]