

2 ה', 1 ת' 3) (נקודות)	<u>מרצה אחראי: פרופ' אלכס ברונשטיין</u> <u>מרצה: פרופ' אבי מנדלסון</u> <u>מרצה ו מתרגל : חיים בסקין</u>	שם הקורס: למידה עמוקה על מאיצים חישוביים	236605
------------------------------	---	---	---------------

סמסטר אביב תשע"ט

שם עברי מקוצר: למידה עמוקה על מאיצים חישוביים

שפת לימוד הקורס: אנגלית

דרישות קדם: תכן לוגי, מת"מ, הסתברות.

סילבוס בעברית

אלגוריתמי למידה עמוקה משמשים כיום לפתרון בעיות בתחומים רבים ומגוונים, החל ממערכות דלות אנרגיה ובעלות יכולות חישוב נמוכות וכלה ב"מערכות על" המתבססות על מחשבי על. על מנת לאפשר פיתוח מהיר של יישומים בתחום, הוצאו לאחרונה מספר סביבות תוכנה חדשות המבוססות על חומרות ייחודיות המאפשרות ביצוע האלגוריתמים באופן יעיל. בקורס זה נתמקד בקשר בין יעילות האלגוריתמים, איכותם, ואמצעי חומרה/תוכנה מתקדמים. נלמד שיטות תוכנה מאפשרות שימוש במאיצי חומרה, כגון כרטיסים גרפיים (GPU), ונרחיב את הדיון לנושאי המחקר העדכניים בתחום.

נושאי לימוד:

- רקע על אלגוריתמי DNN כגון CNN, RNN, תוך מתן דוגמאות שימוש
- שימוש יעיל בחבילת התוכנה PyTorch המיועדת לתמוך בסביבות למידה
- אלגוריתמי למידה עמודה מתקדמים כגון: Variational, Deep Reinforcement Learning, Autoencoders (VAEs)
- יסודות האצת תוכניות בעזרת תכנות מקבילי – CUDA
- אפלקציות של ראייה ממחושבת ועיבוד שפה טבעית בעזרת למידה עמוקה

מקורות:

- Deep “ Ian Goodfellow and Yoshua Bengio and Aaron Courville
MIT Press book ;”Learning
- Nvidia GPUs – The Engine for Deep Learning
<https://developer.nvidia.com/deep-learning>
- /http://pytorch.org - PyTorch
- Morgan ,”Programming Massively Parallel Processors“, Wen-mei W. Hwu
Kaufmann

תוצאות למידה:

1. להבין את העקרונות של פתרון בעיות בעזרת אלגוריתמי למידה עמוקה
2. לדעת לתכנת כרטיסים גרפיים בעזרת CUDA.
3. לדעת איך להשתמש בחבילות תוכנה PyTorch לבניית DNNs
4. ללמד שיטות אופטימיזציה שונות המאפשרות ביצוע יעיל של בעיות למידה על מערכות מחשב מתקדמות
5. ביצוע עבודה מחקרית בתחום

הרכב הציון:

40% עבודות בית 60% פרויקט גמר

פרויקט הגמר יוכל להשתלב בנושא המחקר של הסטודנט או להיות בסיס לפרויקט המשך / עבודה מכינה לקראת הגדרת נושא מחקר

שם הקורס באנגלית

Deep Learning on Computation Accelerators

The course will be taught in English

English syllabus:

Deep learning is widely used in many market segments ranging from mobile devices to supercomputers. Recently different SW packages as well as special HW accelerators were developed to support deep learning. The course will focus on algorithms, programming languages and new SW/HW interfaces that aim to allow execution of deep learning algorithms in a productive and efficient way.

Learning Outcomes:

At the end of the course, the student will

1. Understand and be able to apply notions in deep learning
2. Know how to program GPUs using CUDA.
3. Know how to effectively use PyTorch SW packages
4. Know how to optimize SW and HW performance in deep neural network applications
5. Perform a small research project using the studied notions and techniques

Grad:

40% Drills and 60% final project

Detailed syllabus is below :

Date	WW	Lecture	Tutorial
25.03.18 Last class before Pesah	1	Introduction to ML and Intro to accelerators. (Alex+Avi)	Setup environment , administration and Intro + Final project definition
08.04.18	2	k-NN, linear classifiers, linear regression, loss functions, optimization, descent methods, stochastic gradient descent (Alex)	Numerical computation in Python: numpy and signal processing (Tensor computing) Intro to PyTorch
15.04.18	3	Neural networks: multi-layered	DNN Architectures:

		perceptrons, backpropagation and intro to CNNs, ResNets (Alex)	AlexNet, VGG, GoogLeNet, ResNet, etc.
22.04.18	4	Training neural networks: art and science. Activation functions, initialization, dropout, batch normalization, update rules, data augmentation and domain transfer learning (Chaim)	Training convolution neural networks in PyTorch
29.04.18	5	Recurrent neural networks: RNN, B-RNN, LSTM, GRU and applications (Alex)	RNNs, LSTM training NLP application
06.05.18	6	Generative Models, Variational Auto Encoders, Generative Adversarial Networks, (advanced topics – adversarial attacks) (Chaim)	VAE, GAN training
13.05.18	7	Reinforcement learning, Deep Q-Learning, Policy Gradients (Chaim)	RL
20.05.18	8	CNNs on irregular data -- graph and manifold data (shapes etc.) (Alex)	Style transfer
27.05.18	9	Introduction to floating and point arithmetic, HW considerations Quantization and distillation of DNNs (Chaim)	Teacher-student model
03.06.18	10	HW for DNN inference (Avi/Guest)	Project consultation
10.06.18	11	Data decomposition vs. functional decomposition Elements of parallel computing- (Avi)	Introduction to CUDA
17.06.18	12	Hardware for DNN training GPU as an accelerator –(Avi)	GPU in PyTorch, performance considerations
24.06.18	13	Theoretical insights: why DL works? (Alex)	TBD-Advanced topic
01.07.18			

Assignments:

HW0: MNIST linear classifier

HW1: Hyper-parameters fine tuning on CNN

HW2: LSTM for autocomplete/sentence prediction

HW3: Face synthesis with VAE/GAN (CelebFaces)

HW4: RL on some simple game (backgammon)

HW5: Style transfer (pre trained DNN)

Projects

1. (Aviv Rosenberg) Cardiac data analysis
2. (Yonathan Maor) Aeroacoustics
3. (Sanketh+Nathan?) Audio track separation?
4. (Yochai+Evgeni) RL+genetic algos for bit allocation
5. Speech synthesis (+quantization?)
6. Speech recognition (+quantization?)
7. Generative models/style transfer for music (real time?)
8. Hardware for audio application (echo cancellation? beamforming?)
9. (???) Information bottleneck?
10. Computational imaging (coded aperture?)
11. Image annotation
12. Domain alignment (AutoDial?)
13. Adversarial attacks on RL
14. Distributed training (Compressed sensing?)
15. Recommender systems (learning on graphs?)