פרויקטים בתוכנה סמסטר אביב תשע"ה

Harnessing the power of Deep Learning

מנחה: Roman Kaplan sromanka@tx.technion.ac.il

The hottest topic in ML is Deep Learning, with many startups and large companies entering the field.
Here are some links to understand the impact of Deep Learning:
http://venturebeat.com/2014/12/20/5-deep-learning-startups-to-follow-in-2015/
http://www.technologyreview.com/news/524026/is-google-cornering-the-market-on-deep-learning/
https://www.youtube.com/watch?v=czLI3oLDe8M
Its uses range from object recognition, speech recognition, natural language processing, fraud detection and the list keeps growing.
In this project we will build a real Deep Learning system!
The exact application is for you to choose.
How it works?

1. You choose the exact application
2. We think how to implement the right system.
3. We (mostly you) choose the programming language based on your preference and libraries availability. It can even be a language which you're not familiar with, in order to broaden your programming horizons.

What we will achieve in the project?

1. Develop the program you want to build
2. Experiment with real deep learning systems
3. Freedom to choose your preferred platform to work and acquire advanced programming skills with state of the art technology
4. **Study the latest and fastest developing field in computing**

Similariy search in peer-to-peer networks

מנחה: נעמה קראוס nkraus@tx.technion.ac.il

Peer-to-peer (P2P) networks are highly scalable computer networks. There's no central management in a P2P network, rather, P2P algorithms are fully decentralized.
P2P networks enable storing and searching for data elements; both are done in a distributed manner.
There are several known P2P architectures such as Chord, Pastry, and CAN.
Simialrity search is the task of searching for similar objects (e.g., users, movies and songs, etc.) based on their features.
In this project, you will simulate a P2P network that supports a decentralized simialrity search.

BigData: similarity search of Wikipedia articles

מנחה: נעמה קראוס nkraus@tx.technion.ac.il

In this project, you will experience with a BigData task.
Given some Wikipedia article, the goal is to find similar articles, based on common content such as text and referenced articles.
A possible application is recommending articles to users.
Since Wikipedia dataset is huge, you will tackle the similarity search challenge with BigData algorithms and

tools. Main algorithms and technologies: Localisty Sensitive Hashing (LSH) similarity search, Big data processing, MapReduce, Hadoop/Spark/GraphLab.

Use Machine Learning for Your Desired Purpose

מנחה: Roman Kaplan sromanka@tx.technion.ac.il

Machine learning applications are emerging with new applications by the day. Some of them might sound crazy, but turn out to become a huge success.
In this project YOU decide which problem we're going to solve.
It can be anything you can think of - predicting football scores, estimating the strength of a Texas Holdem hand, stock forecasting, customized movie recommendations from user ratings, or even ranking web search results using personal preference.

How it works?

1. You come-up with the problem to solve
2. We discuss it to define the problem and think exactly what can be achieved
3. We (mostly you) choose the programming language based on your preference and libraries availability. It can even be a language which you're not familiar with, in order to broaden your programming horizons.

What we will achieve in the project?

1. Develop the program you think is useful
2. Experiment with machine learning algorithms to understand their differences and applications
3. Freedom to choose your preferred platform to work
4. Acquire advanced programming skills with state-of-the-art libraries and technology

Note: Knowledge in machine learning isn't a requirement and the relevant parts will be studied during the process

Use Machine Learning for Your Desired Purpose 2

מנחה: Roman Kaplan sromanka@tx.technion.ac.il

Machine learning applications are emerging with new applications by the day. Some of them might sound crazy, but turn out to become a huge success.
In this project YOU decide which problem we're going to solve.
It can be anything you can think of - predicting football scores, estimating the strength of a Texas Holdem hand, stock forecasting, customized movie recommendations from user ratings, or even ranking web search results using personal preference.

How it works?

1. You come-up with the problem to solve
2. We discuss it to define the problem and think exactly what can be achieved
3. We (mostly you) choose the programming language based on your preference and libraries availability. It can even be a language which you're not familiar with, in order to broaden your programming horizons.

What we will achieve in the project?

1. Develop the program you think is useful
2. Experiment with machine learning algorithms to understand their differences and applications
3. Freedom to choose your preferred platform to work
4. Acquire advanced programming skills with state-of-the-art libraries and technology

Note: Knowledge in machine learning isn't a requirement and the relevant parts will be studied during the process

Use Machine Learning for Your Desired Purpose 3

מנחה: Roman Kaplan sromanka@tx.technion.ac.il

Machine learning applications are emerging with new applications by the day. Some of them might sound crazy, but turn out to become a huge success.
In this project YOU decide which problem we're going to solve.
It can be anything you can think of - predicting football scores, estimating the strength of a Texas Holdem hand, stock forecasting, customized movie recommendations from user ratings, or even ranking web search results using personal preference.

How it works?

1. You come-up with the problem to solve
2. We discuss it to define the problem and think exactly what can be achieved
3. We (mostly you) choose the programming language based on your preference and libraries availability. It can even be a language which you're not familiar with, in order to broaden your programming horizons.

What we will achieve in the project?

1. Develop the program you think is useful
2. Experiment with machine learning algorithms to understand their differences and applications
3. Freedom to choose your preferred platform to work
4. Acquire advanced programming skills with state-of-the-art libraries and technology

Note: Knowledge in machine learning isn't a requirement and the relevant parts will be studied during the process

Dare you go

מנחה: כפיר לב-ארי kfirla@campus.technion.ac.il

בפרוייקט הזה נבנה משחק רשת חברתי שמאפשר לכל שחקן להגדיר משימה (קצובה בזמן - שעות, ימים, שבועות..) ולשאר השחקנים להענות למשימה ע"י פרסום תמונה או וידאו שמוכיחים את ביצוע המשימה.
הזוכה במשימה יהיה זה שיקבל את מירב הקולות משאר השחקנים כאשר למפרסם המשימה יש אחוז גבוה יותר בהצבעה.
לכל שחקן יהיה ניקוד שיקבע לפי מספר המשימות שהוא השתתף בהן, הצליח בהן, פרסם וכו'.
הניקוד יקבע מיקום בדירוג כללי של השחקנים שיהיה נגיש לכולם.
את הפרויקט נכתוב כאפליקציה לאנדרואיד בג'אווה, תוך שימוש בצד שרת של
https://www.parse.com
קדם לפרוייקט - הניסויים באנדרואיד ובג'אווה.

Proximity-based playlist

מנחה: כפיר לב-ארי kfirla@campus.technion.ac.il

אם מעניין אתכם לקרוא על אפשרויות לתקשורת ישירה בין פלטפורמות שונות, אתם מוזמנים לקרוא יותר כאן-
https://allseenalliance.org/
בפרויקט הזה נבנה מערכת מבוזרת שמשתמשיה יוכלו (בעזרת מכשיר האנדרואיד שלהם) לשדר לסביבה את ההעדפות המוזיקאליות שלהם, ובכך להשפיע על רשימות ההשמעה במקומות ציבוריים (דוגמאת פאבים, חנויות וכו'), כאשר באותו מקום ציבורי הנגן מחובר למערכת.
הרעיון במעזרת המבוזרת היא שאין שרת מרכזי אליו המשתמשים פונים, אלא הקשרים נוצרים על סמך קרבה ותקשורת ישירה בין הלקוחות לנגנים.

Impoving Linux Load Balancer

מנחה: נעם שלו noams@tx.technion.ac.il

Current implementation of Linux load balancer does not consider offline cores.
In this project we propose an improvement to the load balancing algorithm, implement it on the latest Linux kernel, and test it on a real system.

Website Promotion SEO

מנחה: אילנה דוד ilana@ee.technion.ac.il

ניתוח קישורים. נושא משמעותי בתחום ה- SEO הוא הקישורים המובילים אל האתר. הפרויקט יעסוק בזיהוי הקישורים המובילים אל האתר וביצוע ניתוח מעמיק שלהם כדי להגיע למסקנות. מסקנות שונות יכולות להיות פרופיל הקישורים של האתר, או "מה הסיכוי שהאתר יפגע באחד מעדכוני גוגל הקרובים" וכדו'.
מנחה: ניר אדר

UML State Machine visualizer

מנחה: עובד יצחק ovedi@mellanox.com

UML supports hierarchical state-machine model based on David Harel's Statechart. The W3C consortium defined an XML-based state-machine model specification named SCXML (State Chart XML In this project we will implement a graphical visualization tool with limited editing capabilities for state-machines that are specified through SCXML.
The implementation environment is at the discretion of the students (e.g. Java, C#, Tcl/tk etc.).

MATLAB Application Grid

מנחה: עובד יצחק ovedi@mellanox.com

MATLAB is a numerical computing environment and a programming language. It is a major working tool for many engineering and researchers around the world. MATLAB is also widely used in our department. Many users use MATLAB for solving computationally and data-intensive problems. In order to help these user, the SIPL lab has created a MAS (MATLAB Application Server) that can be accesses using a simple web interface and runs MATLAB scripts on a strong server. The problem with this solution is that even a very strong server may not have the required computation power. The aim of the proposed project is to build a MAG (MATLAB Application Grid) that will have a similar simple web interface but that will run MATLAB scripts on one of the many regular lab computers that form a grid, thus greatly increasing the throughput of the system. The computation power of each computer in the grid will be wisely used only when the computer is idle. In cooperation with the SIPL lab.
This is a two semesters project.
Previous knowledge: C and MATLAB Supervisor: Oved Itzhak Consultant: Yair Moshe

Debugger data-structure visualization framework

מנחה: עובד יצחק ovedi@mellanox.com

Contemporary Debuggers allow inspecting variable in the debugged program based on symbols generated at compile-time. However, this is limited to the language's built-in constructs such as integers, pointers, structs and classes because these are the entities that the compiler knows about and can provide symbolic information for. It cannot provide information about higher level data-structures such as linked lists, trees and hash tables because these are not built-in language constructs.
The Windows Debuggers (http://msdn.microsoft.com/en-us/library/windows/hardware/ff551063%28v=vs.85%29.aspx) provide a very flexible extensibility interface that allow implementing "visualizers" – generate user-friendly visual representation of complex data-structures. However, implementing visualizers can be quite tedious because the extension code runs in a different process than the data – the extension lives in the debugger, which is a process separate from

the debugged program. In this project we will implement a framework that allows writing the extension using the same syntax as if the raw data live in the same process as the extension. The implementation will be mostly in C++ and at the students discretion a part of it can be implemented in some scripting language (e.g. Python, PERL).

Identify Characters from Google Street View with Julia

מנחה: Roman Kaplan sromanka@tx.technion.ac.il

Google Street View contains lots of unidentified text by computers, understandable only to humans.
In this project we will use characters taken from Google Street View images and identifying them thorough software. It differs from traditional character recognition because all data set contains different character fonts and the background is not the same for all images.
In addition, we will use a new emerging scientific programming language called 'Julia', designed for easy implementation of technical programs. Check it out at: 'http://julialang.org/'.
Julia is a high-level language greatly simplifying the amount and complexity of code needed for development of high-level algorithms and has many easy-to-use libraries for parallelism and machine learning algorithms.
What we will achieve in the project?

1. Use real data to develop algorithm with practical applications
2. Experiment with machine learning algorithms to understand their differences and applications
3. Freedom of creativity to choose your preferred method of solution
4. Broaden your arsenal of programming languages with Julia

Knowledge in machine learning isn't a requirement and the relevant parts will be studied during the process

System 106 on the web

מנחה: יהודה ארקין אדר yehuda@microsoft.com

Creating a map and database based WEB system for the management of a municipality accepting and solving complaints 106 system.
The system will be written using database technologies, JavaScript and REST.