English Syllabus

Course Title: Reliability, Equity, and Reproducibility in Modern Machine Learning

Course Code: 236802

Instructor: Avi Tishby

Course Dates:

Semester: Spring 2019

Course Website:
https://moodle.technion.ac.il/course/view.php?id=7998

Prerequisites:
1. 044202 (or equivalent course)
2. 094412 (or equivalent course)
3. 236756 (or equivalent course)
4. 234125 (or equivalent course)

Course Description:

This course focuses on understanding the reliability, fairness, and reproducibility of modern machine learning systems, especially complex systems such as deep learning. We will expose students to the challenges posed by modern learning algorithms and familiarize them with practical methods to address these issues. The course will be taught in a seminar format and will include readings from the forefront of data science. Focus will be placed on evaluating uncertainty, such as gender bias and prevention of gender bias in systems, as well as advanced methods for hypothesis testing across multiple scientific databases and interpretation of learning algorithms.

Outcomes of Learning:

1. Exposure to the challenges of using systems in applications with high risk.
2. Learning methodologies to ensure the reliability of the conclusions obtained from learning systems.

English Syllabus

Course Title: Reliability, Equity, and Reproducibility in Modern Machine Learning

Course Code: 236802

Instructor: Avi Tishby

Course Dates:

Semester: Spring 2019

Course Website:
https://moodle.technion.ac.il/course/view.php?id=7998

Prerequisites:
1. 044202 (or equivalent course)
2. 094412 (or equivalent course)
3. 236756 (or equivalent course)
4. 234125 (or equivalent course)

Course Description:

This course focuses on understanding the reliability, fairness, and reproducibility of modern machine learning systems, especially complex systems such as deep learning. We will expose students to the challenges posed by modern learning algorithms and familiarize them with practical methods to address these issues. The course will be taught in a seminar format and will include readings from the forefront of data science. Focus will be placed on evaluating uncertainty, such as gender bias and prevention of gender bias in systems, as well as advanced methods for hypothesis testing across multiple scientific databases and interpretation of learning algorithms.

Outcomes of Learning:

1. Exposure to the challenges of using systems in applications with high risk.
2. Learning methodologies to ensure the reliability of the conclusions obtained from learning systems.

English Syllabus
In this course, we will study methodologies to guarantee the reliability, robustness, equity, and reproducibility of advanced machine learning systems, such as deep neural nets. We will discuss recent concerns about modern machine learning algorithms and will tackle these by introducing flexible tools that are supported by theoretical guarantees. This course will be given as a seminar and will cover influential papers in the field of data science. We will focus on prediction uncertainty estimation as well as mitigation of discrimination against minorities. We will also explore frameworks for multiple hypothesis testing as powerful tools for making new scientific discoveries and for interpreting complex learning systems.

**Learning Outcomes**

1. Familiarize with challenges in using modern machine learning algorithms in high-stakes applications.
2. Learn flexible methodologies to produce data-driven inferences that can be trusted.

**דרישות המックス**

הציון יבסס על הצגת מאמרים והשתתפות פעילה.

**רשימת ספרות**

מאמרים שעוסקים במциально מובילים בחוגים של למידה חישובית וסטטיסטיקה: