MACHINE LEARNING FOR JUST ABOUT EVERYONE

Expanded offerings for one of MIT's hottest topics reach a broad cross-section of the student population.

By Eric Smalley | Connector Contributor

Introduction to Machine Learning (6.036) meets in MIT's largest lecture hall — and professors still had to whittle down the number of students who enrolled for Spring 2017.

On the first day of Introduction to Machine Learning (6.036) in February 2017, the setting looked more like a sold-out movie theater than an MIT classroom.

At least 700 students converged on the Institute's largest lecture hall, filling every seat and overflowing into another room. Even after being pared to a more manageable crowd of about 550, 6.036 had higher enrollment than any other introductory course in the Department of Electrical Engineering and Computer Science (EECS) — even larger than Introduction to EECS (6.01).

The surge in interest in machine-learning courses should come as no surprise: It mirrors machine learning's transformation from a niche technology to a mainstream area of expertise that's very much in demand. In April 2017, an informal search of Boston-area jobs on Indeed.com, a popular employment website, turned up 805 positions requiring machine-learning experience, more than those seeking candidates with skills in PHP, Perl or robotics. At the same time, Amazon had posted nearly 2,900 jobs companywide for candidates with machine-learning expertise.

Machine learning is an algorithmic approach to data processing that builds models from samples of data sets to describe the data and make predictions accordingly. It's a core component of technologies such as computer vision, natural language processing and robotics. The technology is used in a growing number of fields, including automation, autonomous vehicles, education, finance, health care, marketing, politics, and science.

Machine-learning technologies such as neural networks have been in use for decades, but have become widespread only in recent years because powerful and affordable computing resources and large amounts of data are more available, says Tommi Jaakkola, the Thomas Siebel Professor in EECS and the Institute for Data, Systems, and Society (IDSS). He likens the rise of machine learning to the rise of electricity: "It creates capabilities that weren't there before."

EECS has had a graduate-level intro course, Machine Learning (6.867) for at least 16 years, says Jaakkola, who originally created that course. It's been further developed by its current instructors, Leslie Kaelbling, the Panasonic Professor of Computer Science and Engineering in EECS, and Devavrat Shah, a professor of EECS. As the field took off, so did interest in 6.867. The class began to include a broader range of students, including some from outside EECS and others who were more interested in the technology's applications than the theory behind it. "It was a very mixed population with different demands, so at some point it became untenable to maintain that," Jaakkola says. "We thought we should really create an undergraduate entry-level machine-learning course."

Thus was born 6.036, co-developed and co-taught by Jaakkola and Regina Barzilay, the Delta Electronics Professor of EECS. To address the needs of graduate students who are more

interested in applied machine learning, especially those from outside EECS, the department also created a parallel graduate course, Applied Machine Learning (6.862). Graduate students in 6.862 attend the 6.036 lectures and do the 6.036 assignments, but also undertake a semester-long project supervised by Stefanie Jegelka, X-Consortium Career Development Assistant Professor of EECS.

The sheer number of students interested in the machine-learning courses reflects both the technology's widespread adoption and the high level of expertise it demands. Successful practitioners need to understand how to phrase problems as machine-learning problems, know what methods exist, and be able to choose the appropriate method for each problem, Jegelka says.

The course's popularity reflects machine learning's evolution from a niche area of expertise to one that's in high demand among employers.

The growing use of machine-learning technologies has naturally led to an increase in the number of student internships requiring machine-learning skills. The machine-learning courses emphasize hands-on, applied learning — including 6.867, although that course has a stronger theoretical component than the other two. The hands-on approach not only helps students learn the material, but also gives them the skills they need to land and succeed in internships.

Students in 6.867 work with a large data set, design and run algorithms, modify the algorithms, process the data, and see how the different modifications lead to different types of answers, Shah says. This experience helps students outside the classroom; if they find themselves unsure of which method to use for a particular application, they can start with algorithms from the machine-learning class. "One of the strengths of this course is helping students get ready to do something real," Shah says.

Similarly, 6.036/6.862 teaches students how various machine-learning methods do and don't work in practice and what issues are related to them, Jaakkola says: "They actually have to code up an algorithm and run it, and investigate results on real types of problems."

"The surge in interest in machine-learning courses should come as no surprise: It mirrors machine learning's transformation from a niche technology to a mainstream area of expertise that's very much in demand."

The courses are also helping students outside the department with their research in their fields. Students in 6.862 come from throughout MIT, including from the Departments of Aeronautics and Astronautics, Architecture, Brain and Cognitive Sciences, Chemical Engineering, Civil Engineering, Economics, Mathematics, Mechanical Engineering, and Physics. "These students are defining a project that uses machine learning for their research, so they are working with data from their domain," Jegelka says. "They learn and experience how to formulate the problem, and which methods may work for it, and sometimes also see where inventions from the machine learning side are needed to capture the problem fully."

Projects by students in 6.862 involve a vast range of data-driven topics. Examples include:

• Predicting the energy usage of buildings based on building features.

• Making predictions about molecules, including thermodynamic properties, and the function of proteins based on their 3D structures.

• Addressing problems related to autonomous driving, such as learning driving maneuvers from simulation and recognizing the road from sensor data.

• Identifying cells and membranes from brain imaging.

• Analyzing health insurance uptake in developing countries.

• Detecting "fake news."

• Analyzing transportation policies in Chinese cities.

Ultimately, machine-learning classes are enjoyable for instructors as well, Jegelka says: "It's a lot of fun and very interesting to work with students on such a diverse set of problems and data."