Machine-learning technology powers many aspects of modern society: from web searches to content filtering on social net works to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smartphones. Perhaps more surprisingly, deep learning has produced extremely promising results for various tasks in natural language understanding[14], particularly topic classification, sentiment analysis, question answering[15] and lan guage translation[16,17]. For decades, con structing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a fea ture extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input.