Overview: Information theory is the science of processing, transmitting, storing, and using information. Pioneered by Claude Shannon in 1948 for problems in data compression and reliable communication, it is now relevant to a wide range of fields, including machine learning, statistics, and neuroscience. This course introduces measures of information and uncertainty – such as entropy, mutual information, relative entropy, and differential entropy – and shows how they are connected to practical problems in communication, compression, and inference. Specific topics include lossless data compression, channel capacity, Gaussian channels, and rate distortion theory. The course is appropriate for beginning graduate students in electrical engineering, computer science, statistics, and mathematics who have a good background in probability and linear algebra.

Prerequisites: This course assumes knowledge of probability at the undergraduate level as well as the basics of undergraduate linear algebra and signals and systems. In particular, students should be familiar with the notions of joint and conditional probability distributions, conditional expectation, independence, Gaussian random vectors and processes, and the sampling theorem for bandlimited signals. If you have not already taken these classes (e.g. ECE 581, ECE 555 or their equivalent) or have any doubts about your understanding of these concepts, please discuss them with the instructor.